Universidade de Aveiro Departamento de Biologia
2011

Ana Sofia Paulo Varanda

Aberrant protein synthesis in human HEK293FT cells

Síntese proteíca aberrante em células humanas HEK293FT

Universidade de Aveiro Departamento de Biologia
2011

Ana Sofia Paulo Varanda

Aberrant protein synthesis in human HEK293FT cells

Síntese proteíca aberrante em células humanas HEK293FT

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biologia Aplicada, Ramo Biologia Molecular Celular, realizada sob a orientação científica do Doutor Manuel António da Silva Santos, Professor Associado do Departamento de Biologia da Universidade de Aveiro e co-orientação da Doutora Laura Cristina da Silva Carreto, Investigadora do Departamento de Biologia da Universidade de Aveiro.

o júri

presidente	Doutora Maria Adelaide de Pinho Almeida Professora Auxiliar da Universidade de Aveiro
arguente	Doutora Margarida Alexandra Vaz Caldeira Investigadora Auxiliar do Centro de Neurociências e Biologia Celular da Universidade de Coimbra
orientador	Doutor Manuel António da Silva Santos Professor associado da Universidade de Aveiro
co-orientador	Doutora Laura Cristina da Silva Carreto Investigadora da Universidade de Aveiro

Agradecimentos

Em primeiro lugar gostaria de agradecer ao meu orientador Doutor Manuel Santos por ter acreditado em mim, por todos os ensinamentos e acompanhamento ao longo deste ano.

Um agradecimento especial ao Doutor Carlos Duarte pela forma como me acolheu no Centro de Neurociências e Biologia Celular em Coimbra, se interessou pelo meu trabalho, pelos conselhos e conhecimentos que me transmitiu.

Aos grupos do Doutor Carlos Duarte e Doutora Ana Luísa Carvalho do CNC agradeço a alegria, bom ambiente, paciência e todo o tempo que disponibilizaram para me ensinar, em especial à Marta, Sandra, Graciano, Margarida e Miranda.

Ao Doutor Artur Paiva e ao Tiago Carvalheiro agradeço a colaboração no meu projecto e a disponibilidade com que sempre me acolheram no Centro de Histocompatibilidade em Coimbra.

A todo o grupo do laboratório de Biologia do RNA de Aveiro, agradeço os conselhos, discussões frutíferas e bons momentos, em especial à Doutora Laura Carreto e aos meus colegas de Mestrado.

Aos Biólogos e Biomédicos agradeço a força, amizade e a alegre companhia em toda esta caminhada, em especial à Cátia, Inês e Tiago.

A todos os meus amigos, em especial à Ana, um muito obrigada pela presença calorosa e constante em tudo o que faço, ao longo de todos estes anos.

Ao João, pelo companheirismo, carinho, amizade e compreensão em todos os bons e maus momentos.

Aos meus pais, por serem a “rede” por baixo do meu trapézio, pelos ensinamentos, apoio incondicional e claro por me terem transmitido os “genes” do gosto pela ciência, um grande beijinho e muito obrigada por tornarem este trabalho possível.

palavras-chave

Incorporação de erros durante a síntese proteica; proteínas aberrantes; controlo de qualidade de proteínas; resposta ao stress; alterações transcripcionais

A síntese proteica é um mecanismo sujeito a uma apertada regulação, pois a manutenção de funções celulares essenciais está dependente da fidelidade deste processo. Em condições fisiológicas normais podem ocorrer erros, a uma frequência de aproximadamente 10^{-4} erros por codão descodificado. Este nível de erro é tolerado pelas células, mas quando a frequência destes aumenta, os mecanismos de controlo de qualidade das proteínas podem falhar levando à acumulação de proteínas aberrantes (misfolded) que tendem a formar agregados tóxicos, podem comprometer a viabilidade celular, alterar a homeostasia e levar ao desenvolvimento de doenças.

Com o objectivo de elucidar os mecanismos de resposta à acumulação de erros durante a síntese proteica e entender como o aumento de proteínas aberrantes pode levar ao desenvolvimento de doenças e degeneração celular, foram utilizadas células HEK293FT expostas a canavanina e azetidine-2-carboxylic acid (AZC), análogos da arginina e prolina respectivamente. A incorporação destes análogos de aminoácidos leva a uma síntese proteica aberrante e origina proteínas que tendem a agregar e ter consequências tóxicas para as células. No nosso estudo observámos que a incorporação de análogos de aminoácidos levou a uma diminuição ligeira da viabilidade celular, assim como levou ao aumento do número de células nas fases G2/M e S do ciclo celular. Também foi detectado um aumento de proteínas conjugadas com ubiquitina, não se observando alteração na actividade do proteasoma. Isto pode indicar que a este nível de erro a que as células estão sujeitas, os mecanismos de controlo de qualidade de proteínas estão a ser activados de forma a evitarem a agregação proteica. Através das análises ao transcriptoma observou-se que os genes cuja expressão diminuiu estão relacionados com a matriz extracelular e adesão celular e os genes cuja expressão aumentou, estão envolvidos na regulação negativa da transcrição e na resposta a proteínas aberrantes.

Este estudo permitiu adquirir novos conhecimentos acerca da resposta celular à incorporação de erros durante a síntese proteica, através da análise de alterações transcripcionais causadas pelo stress proteotóxico. As células HEK293FT são um bom modelo de estudo para compreender as bases moleculares de doenças humanas originadas por uma síntese proteica aberrante.

keywords

mRNA mistranslation; protein misfolding; protein quality control; stress response; transcription alterations

abstract

Protein synthesis is tightly regulated and fidelity in this process is essential for maintenance of essential cellular functions. Under normal physiological conditions errors can occur, at frequencies around 10^{-4} errors per codon decoded. This level of mistranslation can be tolerated by cells. When the frequencies of errors increase mechanisms of protein quality control can fail leading to accumulation of misfolded proteins, which are more prone to form toxic aggregates, can compromise cell viability, disrupt cellular homeostasis and lead to the development of diseases.

In order to elucidate how cells respond to mistranslation and understand how protein accumulation can cause disease and cell degeneration, we exposed human HEK293FT cells to canavanine and azetidine-2-carboxylic acid (AZC) which are analogues of arginine and proline, respectively. Their misincorporation into proteins leads to erroneous protein synthesis, misfolding and likely to protein aggregation. Such mistranslation event decreased slightly cellular viability and increased the number of cells arrested in G2/M and S phases of cell cycle. In HEK293FT cells was detected an increase in proteins conjugated with ubiquitin, without altering proteasome activity, which may indicate that at this level of mistranslation, protein quality control mechanisms are active to counteract the formation of protein aggregates. Transcriptional analysis showed that down-regulated genes were mainly associated with extracellular matrix and cell adhesion and up-regulated genes were involved in negative regulation of transcription and response to unfolded proteins.

This study provides new insights into the response of human cells to mistranslation by giving a global analysis of transcriptional alterations that occur in response to proteotoxic stress. HEK293FT cells can be a good model to understand the molecular basis of human diseases caused by mRNA mistranslation.

Table of contents

List of abbreviations

aa	amino acid
aa-AMP	aminoacyl-adenylate
aaRS	aminoacyl-tRNA synthetase
aa-tRNA	amino acid-tRNA
ADLP	*adrenoleukodystrophy*
ADP	adenosine diphosphate
AFG3L1P	*AGF3 ATPase family gene 3-like 1 (S. cerevisiae)*
AHSA1	*AHA1, activator of heat shock 90kDa protein ATPase homolog 1 (yeast)*
Ala	alanine
ANO7	*anoctamin 7*
Arg	arginine
ATP	adenosine 5'-triphosphate
ATPase	adenosine triphosphatase
AZC	azetidine-2-carboxylic acid
BAD	*BCL2-associated agonist of cell deah*
BBS10	*Bardet-Biedl syndorme 10*
BBS12	*Bardet-Biedl syndorme 12*
BBS6	*Bardet-Biedl syndorme 6*
CD81	*CD81 molecule*
CDKN1A	*cyclin-dependent kinase inhibitor 1A*
CIDEC	*cell death-inducing DFFA-like effector c*
COL13A1	*collagen, type XIII, alpha 1*
COL2A1	*collagen, type II, alpha 1*
COL5A2	*collagen, type V, alpha 2*
CORO2A	*coronin, actin binding protein, 2A*
cRNA	complementary ribonucleic acid
DDIT3	*DNA-damage-inducible transcript 3*
DDIT4	*DNA-damage-inducible transcript 4*
DHRS11	*Dehydrogenase/reductase (SDR family) member 11*
DMEM	Dulbecco's modified Eagle´s medium
DNA	deoxyribonucleic acid
DNAJB1	*DnaJ(Hsp40) homolog, subfamily B, member 1*
DUBs	deubiquitinating enzymes
eEFs	eukaryotic elongation factors
EGFLAM	*EGF-like, fibronectin type III and laminin G domains*
eIFs	eukaryotic initiation factors
ELISA	enzyme-linked immunosorbent assay
ER	endoplasmic reticulum

ERAD	endoplasmic reticulum associated degradation
eRFs	eukaryotic release factors
FBS	Fetal bovine serum
FIBP	*fibroblast growth factor (acidic) intracellular binding protein*
FUT11	*fucosyltransferase 11(alpha (1,3) fucosyltransferase)*
GDP	guanosine 5'-diphosphate
Gly	glycine
GTP	guanosine 5'-triphosphate
GTPase	guanosine triphosphatase
GUSBP1	*glucuronidase, beta pseudogene 1*
HOXA4	*homeobox A4*
Hsp40	heat shock protein 40
Hsp70	heat shock protein 70
Hsp90	heat shock protein 90
HSPH1	*heat shock 105kDa/110kDa protein 1*
Ile	isoleucine
KAZALD1	*Kazal-type serine peptidase inhibitor domain 1*
LAMA1	*laminin, alpha 1*
LAMA5	*laminin, alpha 5*
LAMC2	*laminin, gamma 2*
Leu	leucine
Lys	lysine
MAX	*MYC associated factor X*
Met	methionine
MKKS	*Mckusick-Kaufman syndrome*
mRNA	messenger ribonucleic acid
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide
NADH	nicotiamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NES	*nestin*
OBSL1	*obscurin-like 1*
ORF	open reading frame
PAB	poly A binding protein
PBS	phosphate buffered saline
PDLIM7	*PDZ and LIM domain 7 (enigma)*
PEX26	*peroxisomal biogenesis factor 26*
PIGL	*phosphatidylinositol glycan anchor biosynthesis, class L*
PPi	inorganic pyrophosphate
PRODH	*proline dehydrogenase (oxidase) 1*
PRRT2	*Proline-rich transmembrane protein 2*
PVDF	polyvinylidene fluoride

RAB26	*RAB26, member RAS oncogene family*
RNA	ribonucleic acid
ROS	reactive oxygen species
RPRML	*reprimo-like*
rRNA	ribosomal ribonucleic acid
SAP30BP	*SAPP30 binding protein*
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
SECTM1	*secreted and transmembrane 1*
Ser	serine
SETD4	*SET domain containing 4*
SLC25A15	*solute carrierfamily 25 (mitochondrial carrier; ornithine transporter) member15*
SLITRK6	*SLIT and NTRK-like family, member 6*
SMPD3	*sphingomyelin phosphodiesterase 3, neutral membrane (neutral sphingomyelinase II)*
SPON2	*spondin 2, extracellular matrix protein*
SRGAP2	*SLIT-ROBO Rho GTPase activating protein 2*
STRA	*stimulated by retinoic acid gene 6 homolog (mouse)*
TCEA2	*transcription elongation factor a (SII), 2*
TCOF1	*Treacher Collins-Franceschetty syndrome 1*
TMEM107	*transmembrane protein 107*
TMEM129	*transmembrane protein 129*
TMEM80	*transmembrane protein 80*
tRNA	transfer ribonucleic acid
TTC39A	*tetratricopeptide repeat domain 39A*
Tyr	tyrosine
UBD	ubiquitin binding domains
UPR	unfolded protein response
UPS	ubiquitin-proteasome system
Val	valine
ZNF320	*zinc finger protein 320*

Chapter I – Introduction

1.1 The Process of Translation

The genetic information of all the organisms is stored in the genome. This information is used to produce proteins at the appropriate time and place (Lodish, Berk et al. 2005). The DNA codifies messenger ribonucleic acid (mRNA) through a process called transcription. The mRNAs carry the information that cells use to assemble chains of amino acids during the process of protein synthesis, which is carried out by the ribosome that interprets mRNA information with the aid of transfer ribonucleic acid (tRNA) and translation factors (Lewin 2004). Therefore, ribosomes are important macromolecular machines that translate the mRNA code. These machines are composed of ribosomal ribonucleic acid (rRNA) and proteins. In eukaryotes, ribosomes are composed by a small subunit (40S) and a large subunit (60S). In the small subunit (40S) there are three tRNA binding sites: aminoacyl site (A site), peptidyl site (P site) and exit site (E site) (Lodish, Berk et al. 2005).

Translation can be divided into three main steps: Initiation, Elongation and Termination. During initiation the 40S subunit joins the mRNA and searches for the initiation codon, using the initiator tRNA ($tRNA_i$) as the searching tool. During elongation aminoacyl tRNAs enter in the A site of the ribosome, anticodons of acylated tRNA are complementary to codons in the mRNA, so decoding takes place in the ribosomal A site. The ribosome then translocates through the mRNA and catalyzes peptide bonds between amino acids, forming a polypeptide chain. Termination occurs when the ribosome encounters a stop codon, which leads to dissociation of the polypeptide chain from the ribosome. The ribosomal subunits are dissociated, mRNA and deacylated tRNA are released and the components can be recycled for initiation of another round of translation (Figure 1). These translational processes are conserved among of living organisms, but there are some differences between bacteria and eukaryotes (Kapp and Lorsch 2004; Lodish, Berk et al. 2005). For the purpose of my work I will focus on the process of eukaryotic translation.

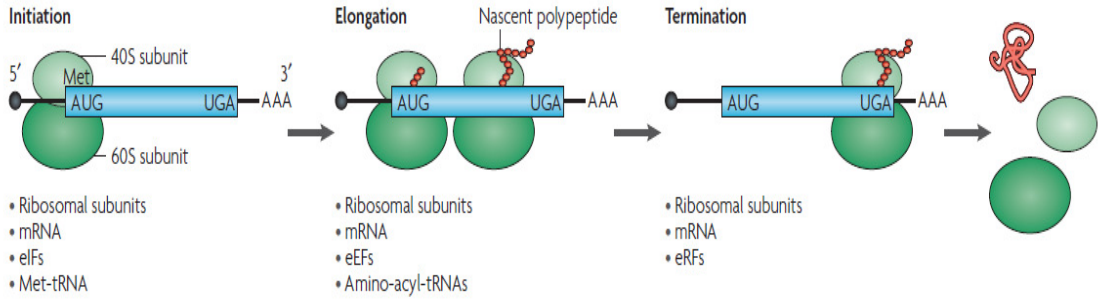


Figure 1. The process of translation in eukaryotes showing the three main steps: Initiation, Elongation and Termination. In blue is represented the open reading frame (ORF); eIFs – eukaryotic initiation factors; eEFs – eukaryotic elongation factors; eRFs – eukaryotic release factors; Met – methionine [adapted from (Scheper, Knaap et al. 2007)].

1.1.1 Initiation

The first step of eukaryotic initiation is the assembly of the ternary complex eIF2.GTP.Met-tRNA$_i$. The eIF2 binds GTP which is hydrolyzed to GDP during initiation. Because eIF2 has high affinity for GDP, another factor eIF2B is needed to recycle GDP to GTP. The ternary complex binds to the small subunit 40S with the help of eIF1, eIF1A and eIF3 originating the 43S complex (Kapp and Lorsch 2004).

The 5'cap of mRNA is recognized by the eIF4F complex, which contains the protein eIF4eE that recognizes the terminal 7-methylguanosine of the mRNA 5'end. The eIF4B and eIF4H (RNA binding proteins, the second recently discovered in mammalian systems) bind to mRNA and unwind secondary structure in the 5'-untranslated region (5'UTR) using the ATP dependent action of the eIF4A subunit an RNA-helicase. The eIF4E and eIF4A bind to eIF4G, which is thought to act as an adapter protein. Together eIFs 4A, 4E and 4G form the eIF4F complex (Kapp and Lorsch 2004).

The eIF3 which is a giant heteromultimeric complex, and the poly A binding protein, which recognizes the 3'-poly A tail, help the assembly of the mRNA in the 43S complex. The eIF3 serves as a scaffold to alter the conformation of the 40S subunit, allowing easier access for mRNA, while PAB interacts with eIF4G leading to circularization of the mRNA, which is thought to stimulate translation by promoting the binding of the 43S complex to the mRNA. This complex scans the mRNA from 5'-3' direction, and stops when the Met-tRNA$_i$ anticodon recognizes the initiation codon in the mRNA. The scanning process is thought to require ATP hydrolysis, although the ATPase

involved has not yet been identified. eIF1 and eIF1A also participate in the scanning process. The selection of the initiation codon (AUG) is facilitated by specific surrounding nucleotides which form a favorable sequence context (Kapp and Lorsch 2004; Lodish, Berk et al. 2005). The best context contains a purine at position -3 and a G a position +4. Then occurs the hydrolysis of GTP bound to eIF2, a reaction facilitated by eIF5 a GTPase-activating protein (GAP). eIF2.GDP releases the Met-tRNA$_i$ in the P-site of the small subunit (40S) and dissociates from the complex. Some other factors, like eIFs 1, 1A, 3 and 5, also dissociate at this stage. eIF5B.GTP binds to the complex probably after the release of eIF2.GDP. The large ribosomal subunit 60S joins to 40S.Met-tRNA$_i$.mRNA complex and this requires hydrolysis of GTP bound to eIF5B. The eIF5B stabilizes tRNA$_i$ and facilitates the joining of the two ribosomal subunits. This step is irreversible since the ribosomal subunits do not dissociate until the entire mRNA is translated and proteins synthesis is terminated. eIF5B.GDP has low affinity for the ribosome so it dissociates from the complex (Figure 2) (Kapp and Lorsch 2004).

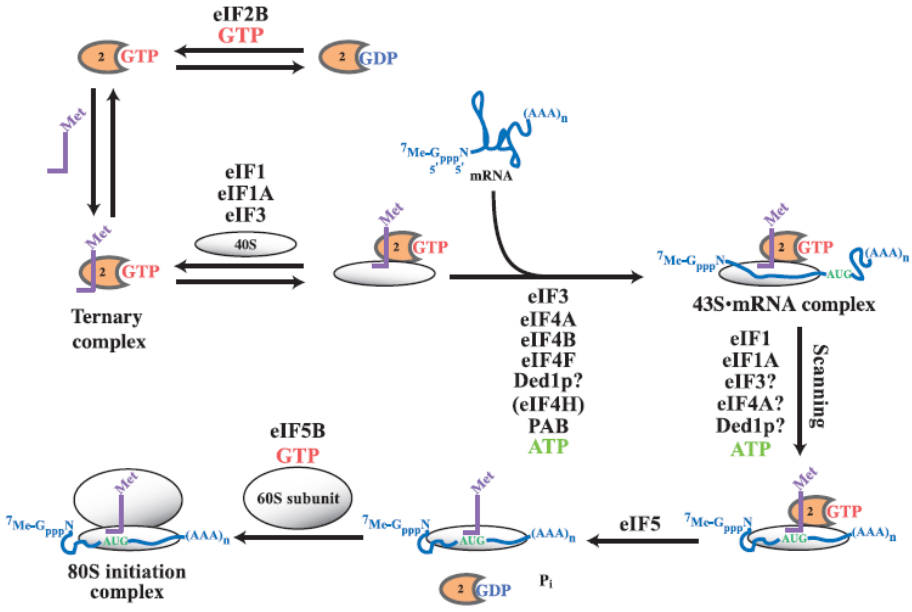


Figure 2. Scheme of the first step in Eukaryotic Translation – Initiation. The involvement of some of the initiation factors indicated is still controversial, namely Ded1p in the assembly of 43S.mRNA complex, or eIF3 in scanning [adapted from (Kapp and Lorsch 2004)].

1.1.2 Elongation

Elongation starts with a peptidyl-tRNA in the P site, and a vacant A site which accepts aminoacyl-tRNA carried in the form of a ternary complex with eEF1A.GTP (Kapp and Lorsch 2004). Only the cognate tRNA carrying the correct amino acid, can enter the elongation phase because correct anticodon-codon interactions are required for the releasing of tRNAs at the A site. Binding of cognate tRNA to the A site triggers ATP hydrolysis which promotes a conformational change in the ribosome that leads to tight binding of the aminoacyl-tRNA in the A site and the release of the eEF1A.GDP complex. The conformational change in the ribosome also positions the 3' end of aminoacylated tRNA in the A site in proximity to the 3' end of peptidyl-tRNA in the P site (Lodish, Berk et al. 2005). The ribosomal peptidyl transferase center then catalyzes the formation of the peptide bond between incoming amino acid and peptidyl-tRNA. This results in the formation of deacylated tRNA which moves to the E site during ribosome translocation (Kapp and Lorsch 2004). This step is accomplished by eEF2 which hydrolysis GTP and facilitates translocation (Lodish, Berk et al. 2005). After translocation the deacylated tRNA is located in the E site and the peptidyl-tRNA in the P site. The A site is free to receive another aminoacyl-tRNA complexed with eEF1A.GTP and this cycle is repeated untill a stop codon arrives at the A site, which activates the process of termination (Figure 3). After the hydrolysis of GTP and the release of aminoacyl tRNA, into the ribosome eEF1A.GDP is released and is recycled to its GTP-bound form to participate in successive rounds of polypeptide elongation. eEF1B, a multifactor complex catalyses this exchange of GDP for GTP (Kapp and Lorsch 2004).

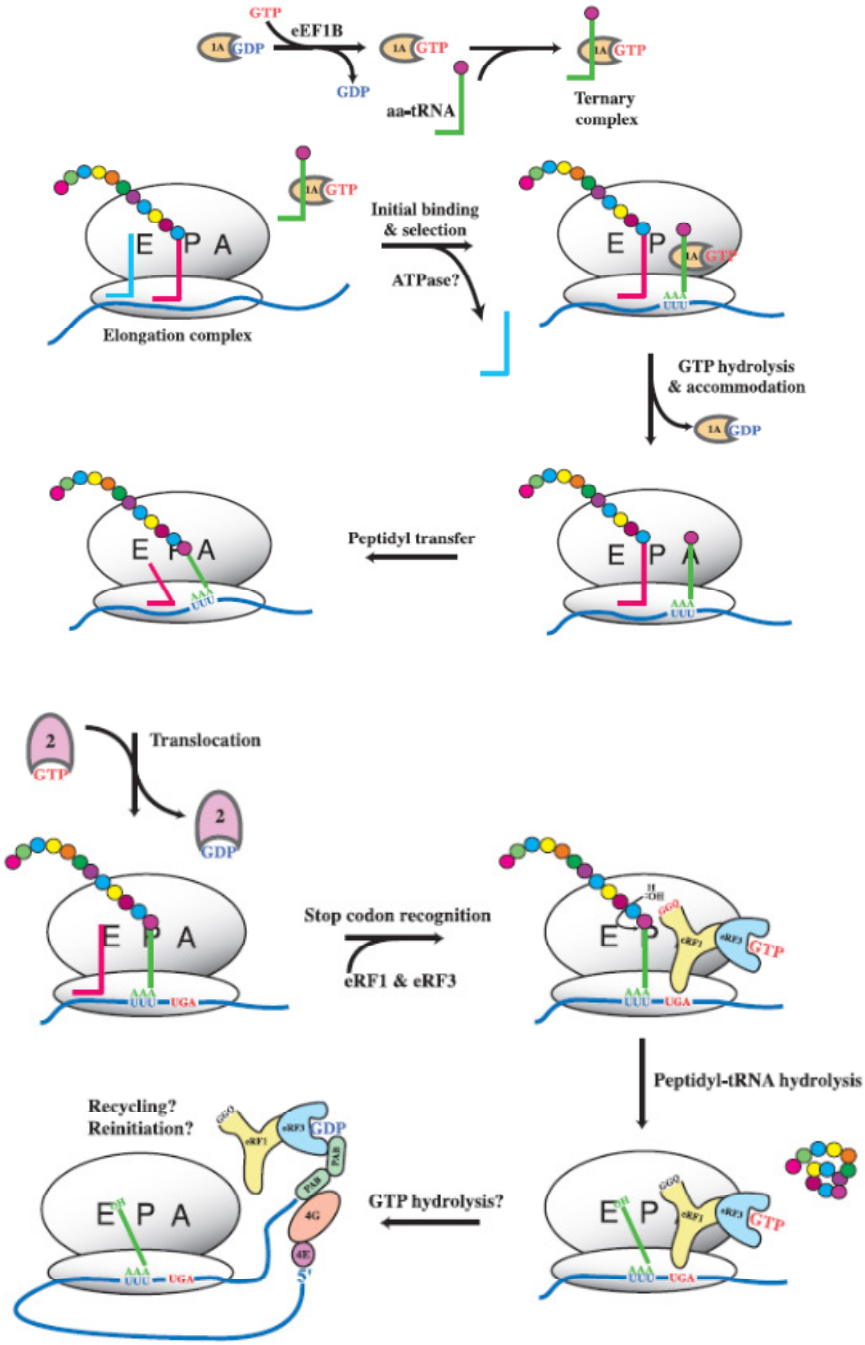


Figure 3. Translation process. Scheme of Elongation, Termination and Recycling in Eukaryotes (Kapp and Lorsch 2004).

1.1.3 Termination and Recycling

The presence of a stop codon in the ribosome A site activates the hydrolysis of the ester bond that links the polypeptide chain to the P site_tRNA. Hydrolysis is catalyzed by the peptidyl transferase center of the ribosome, in response to the activity of class 1 release factors. This class of factors has similar structure to that of tRNAs and bind to the ribosomal A site, recognizing stop codons directly. The eRF1 has an omnipotent decoding capacity as it can promote the hydrolysis of peptidyl-tRNA in response to any of the three stop codons (UAA, UAG, UGA). Class 2 release factors are GTPases that stimulate the activity of class 1 release factors regardless of which stop codon the class 1 factor has engaged. In eukaryotes the class 2 release factors is eRF3 whose activity promotes the release of eRF1 from the ribosome following peptidyl-tRNA hydrolysis (Figure 3) (Kapp and Lorsch 2004; Lodish, Berk et al. 2005).

The last stage of translation is the recycling of the various components of the ribosome to use them in another round of initiation. In eukaryotes this step remains obscure, since little is known about how the ribosomal subunits are dissociated and how mRNA and deacylated tRNA are released. Some studies showed that eIF3 can have an anti association activity due to induction of a conformational change in the 40S subunit, increasing the rate of subunits dissociation as well as lowering the rate of association. The close loop model of eukaryotic mRNAs suggests that the 40S subunit is not released. Instead, it may be shuttled across or over the poly (A) tail back to the 5’ end of the mRNA. This close loop facilitates the reinitiation of translation (Scheper, Knaap et al. 2007).

1.2. Genetic code

The genetic code was established in 1960's, and establishes the rules used by cells to translate the DNA information into protein information (Figure 4). This code is based on nucleotide triplets (codons), and these codons do not overlap along the mRNA. The 4 ribonucleotides (adenosine –A, guanine – G, uridine – U and cyclidine – C) can form 64 possible codons. 61 specify individual amino acids and 3 are stop codons, i.e. do not specify amino acids (UAA, UGA and UAG). Some amino acids such as methionine are specified by only one codon, but most amino acids are encoded by more than one codon. For example leucine and serine are specified by six different codons each. These codons that specify the same amino acid are called synonymous codons. The genetic code is therefore degenerated because more than one codon can specify the same amino acid (Agris 2004; Lewin 2004).

Many organisms contain fewer tRNAs than the number of codons. In order for tRNA to recognize more than one codon the first base in the first position of the anticodon must be able to partner alternative bases in the corresponding third position of the codon. According to the Wobble hypothesis, the pairing between codon and anticodon at the first two codon positions follows the usual rules of base pairing, but exceptional wobbles occur at the third position. This explains why multiple codons that specify the same amino acid often differ at the third base position only. This is possible because the conformation of the tRNA anticodon loop allows some flexibility at the first base of the anticodon (Agris 2004; Lewin 2004)

The first codon of mRNAs in eukaryotic cells is AUG and specifies methionine, and the mRNA sequence from the start codon to the stop codon is called open reading frame (ORF), which specifies the precise linear sequence of amino acids that form the polypeptide chain (Lodish, Berk et al. 2005).

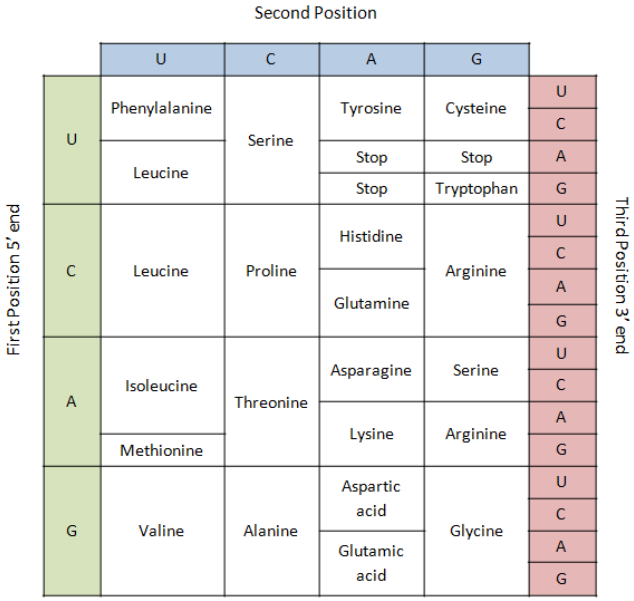


Figure 4. The Universal Genetic Code [adapted from (Lodish, Berk et al. 2005)]

1.2.1 tRNAs

tRNAs are "adapter" molecules that match codons to particular amino acids and are crucial translation factors that define the genetic code. tRNAs have two crucial properties: each tRNA binds a single amino acid (covalently linked) and contains a trinucleotide, the anticodon, which is complementary to the codon in mRNA that is assigned to that amino acid. The anticodon recognizes the codon via complementary base pairing (Lewin 2004; Lodish, Berk et al. 2005).

These molecules evolved from a single ancestor by duplication and subsequent mutations (Lenhard, Orellana et al. 1999). There are about 50-100 tRNAs in animal and plant cells and their genes are transcribed by RNA polymerase III. The number of tRNA differs from the number of amino acids since many amino acids have more than one tRNA to which they can attach and the number of codons in the genetic code since many tRNAs can pair with more than one codon (Lewin 2004; Lodish, Berk et al. 2005; Hopper, Pai et al. 2010).

tRNAs are 74-95 nucleotides long and have common secondary and tertiary structures. The function of tRNA depends on their precise three-dimensional structure. In

solution these molecules can fold into a secondary structure that resembles a cloverleaf (Figure 5) (Lewin 2004; Lodish, Berk et al. 2005).

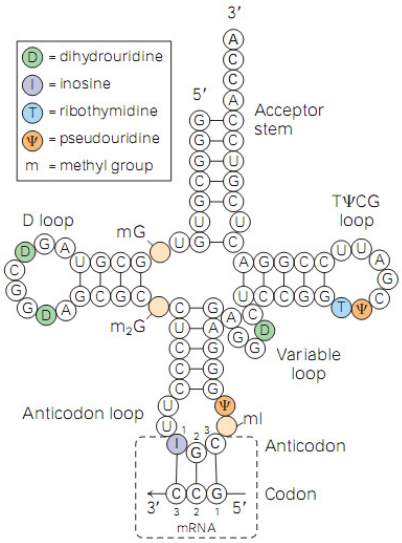


Figure 5. The Alanine tRNA ($tRNA^{Ala}$) from yeast - secondary structure: cloverleaf (Lodish, Berk et al. 2005)

In this structure, complementary base pairing forms stem-loops which are the arms of the tRNA. There are four major arms: the acceptor arm, the TΨC arm, the anticodon arm and the D arm. The acceptor arm consists of a base-paired stem ending in an unpaired sequence with a free at 2'- or 3'-OH. This group can link the amino acid. The TΨC arm has a sequence of timine, pseudouridine (modified base) and cytosine. The anticodon arm contains the anticodon triplet in the center of the loop and the D arm is characterized by the presence of the modified base dihydrouridine. There is also an extra arm located between the TΨC and the anticodon arms whose length can vary between 3 to 21 bases. In all tRNAs the 3'end has the sequence CCA, which is added after synthesis and processing of the tRNAs are complete (Lewin 2004; Nakanishi and Nureki 2005).

tRNAs can be divided in two classes (class I and II) according to the length of the extra arm. Class I tRNAs have a short extra arm, 4-5 nucleotides (almost all tRNAs), while class II tRNAs have long extra arm with at least 11 nucleotides. tRNAs with long extra arms are phylogenetically well-conserved and include Leucine tRNAs ($tRNAs^{Leu}$), Serine tRNAs ($tRNAs^{Ser}$) in eukaryotes and Tyrosine tRNAs ($tRNAs^{Tyr}$), $tRNAs^{Leu}$, $tRNAs^{Ser}$ in

prokaryotes . The long extra arm may be involved in the recognition of these tRNAs by their cognate aminoacyl tRNA synthetases or in discrimination against non-cogante synthetases (Lenhard, Orellana et al. 1999).

Some positions in tRNA are conserved (or invariant). That is, a particular base is always found in a particular position. Others are semiconserved (semi-invariant), i.e. restricted to one chemical type of base (purine or pyrimidine) (Lewin 2004).

In all tRNAs, the cloverleaf structure forms a tertiary L-shaped structure which is created by hydrogen bonding (Figure 6). There are two functional domains in the L-shaped structure which had independent origins, namely the acceptor arm which is at one end of the tRNA and the anticodon arm at the other end of the tRNA L-type structure (Lenhard, Orellana et al. 1999; Ishitani, Yokoyama et al. 2008) The structure of tRNAs is consistent with their role in protein synthesis, for this reason the amino acid (attached to the acceptor stem) and the anticodon are at opposite ends of the tRNA L-type structure (Lewin 2004; Lodish, Berk et al. 2005).

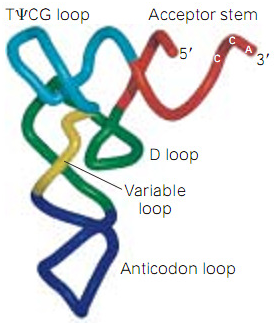


Figure 6. Three-dimentional model of the tRNAs: L-shape (Lodish, Berk et al. 2005)

Aminoacyl-tRNA synthetases (aaRSs) are the enzymes responsible for the attachment of the amino acid to the tRNA. The aaRS recognize their cognate tRNAs trough a limited number of nucleotide residues called identity determinants, which are commonly located in the tRNA anticodon, the acceptor stem and the associated "discriminator" base at position 73. The extra arm may also contain identity elements for tRNA recognition. Each family of tRNAs has a specific discriminator base. For example, all $tRNAs^{Ser}$ have G_{73} and all $tRNAs^{Leu}$ have A_{73} discriminator bases (Asahara, Himeno et al. 1993; Lenhard, Orellana et al. 1999). Recent studies revealed that major identity

determinants are conserved during evolution, but minor identity elements often change (Lenhard, Orellana et al. 1999; Sekine, Nureki et al. 1999). The tRNA recognition system may also involve identity anti-determinants which are responsible for preventing their recognition by non-cognate aaRSs.

Recent discoveries indicate that tRNAs can have additional roles beyond participation in protein synthesis. For example, some tRNAs may be involved in the control of apoptosis (programmed cell death), through binding to cytochrome *c*, a molecule that when released from the mitochondria is responsible for the initiation of the intrinsic apoptotic pathway (Raam and Salvesen 2010). The tRNA molecules can also be precursors of other small non-coding RNAs. Indeed, studies in HeLa cells identified highly abundant small RNA fragments that were derived from mature tRNAs processed by enzymes responsible for formation of small regulatory RNAs (Cole, Sobala et al. 2009; Miyoshi, Miyoshi et al. 2010).

1.2.3 Aminoacyl-tRNA synthetases (aaRSs)

Before translation, amino acids must be attached to their cognate tRNAs by aaRSs. These enzymes establish therefore the genetic code by matching amino acids with tRNA anticodons (Pouplana and Schimmel 2001a; Schimmel 2008). These ancient enzymes are ubiquitously expressed and universally distributed across the phylogenetic tree suggesting that they are among the oldest polypeptide families (Pouplana and Schimmel 2001b; O'Donoghue and Luthey-Schulten 2003); and are intimately tied to the historical development of the genetic code (Pouplana and Schimmel 2001b; O'Donoghue and Luthey-Schulten 2003; Park, Schimmel et al. 2008).

Each aaRS recognizes and activates a single amino acid through a universally conserved two-step mechanism. In the first step (1) the amino acid (aa) is condensed with adenosine triphosphate (ATP) forming a tighly bound aminoacyl adenylate. This reaction is accompanied by the release of inorganic pyrophosphate (PPi). In the second step (2) the activated amino acid is transferred to the 3' end of tRNA and originates aminoacyl-tRNA (aa-tRNA) releasing adenosine monophosphate (AMP) (Figure 7) (Schimmel 2008).

(1) E+ aa +ATP→E(aa-AMP)+PPi

(2) E(aa-AMP)+tRNA→E+ aa-tRNA+AMP

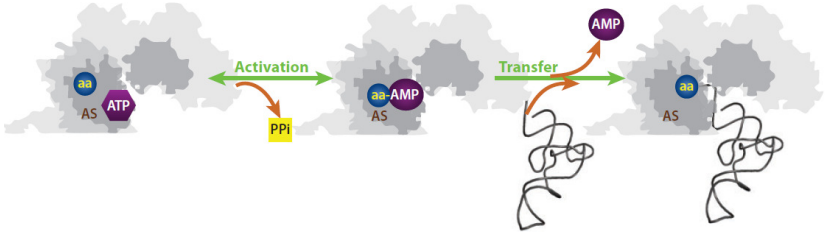


Figure 7. The aminoacylation reaction (Ling, Reynolds et al. 2009)

The aaRSs are divided into two classes (Class I and Class II) which contain 10 enzymes each. This division in two classes is based on differences in the structural topology of their active sites (Eriani, Delarue et al. 1990). Interesting, each subclass originated from an ancient and distinct single-domain protein and there is no evidence for the existence a common ancestor (Pouplana and Schimmel 2001a).

Class I aaRSs are generally monomeric and have in their catalytic domain a Rossman nucleotide binding fold (Schimmel 2008). They approach tRNA molecules from the minor groove of the tRNA acceptor stem and aminoacylate the terminal adenosine (of the CCA-3' terminal acceptor stem) at the 2'-OH position (Sprinzl and Cramer 1975). In contrast Class II aaRSs are typically multimeric enzymes and their active sites contain seven-stranded antiparallel β-fold with flanking α-helices (Schimmel 2008). They approach tRNAs from the major groove and charge the terminal adenosine at the 3'-OH position, based on studies carried out in *E.coli* (Sprinzl and Cramer 1975).

The aaRSs can be also divided into three subclasses within each class, namely Ia,b,c and IIa,b,c. These subclasses represent the enzymes that are more closely related to each other based on their mode of binding to the tRNA acceptor stem (Pouplana and Schimmel 2001b). Subclasses Ia and IIa catalyze aminoacylation reactions of many of the hydrophobic amino acids, while subclasses Ib and IIb capture the carboxyl side-chain amino acids and the amidated (NH_2) derivates. Subclasses Ic and IIc catalyze aminoacylation reactions for aromatic amino acids (Figure 8) (Schimmel 2008).

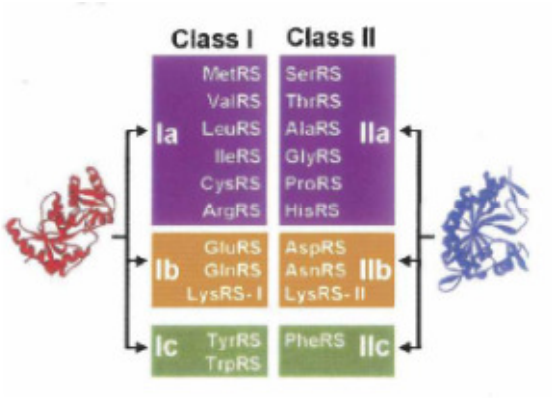


Figure 8. Classes and subclasses of aaRS. LysRS are found in class II, but there are also examples of class I LysRS (Schimmel 2008).

To ensure accurate aminoacylation, these enzymes have specific determinants that are crucial to recognize the cognate tRNA, like Leu-136 in Glutaminyl-tRNA synthetase (GlnRS) of *E. coli*, which is important for tRNA specificity of $tRNA^{Gln}$. The acceptor stem of the tRNA and the anticodon sequence are crucial to recognition of the appropriate tRNA by this enzyme (Sherman and Söll 1996).

The active site of aaRSs (during the activation step) preferentially selects the cognate amino acid due to differences in the side chain binding energies of the substrate (Favorova 1984). Usually the innate specificity for recognition of cognate amino acid-tRNA pair is often sufficient to ensure correct aminoacylation, but sometimes, when the cognate amino acid displays high structural similarity to other non cognate amino acid (for example amino acids that differ by a single methyl group), it is difficult for aaRS to distinguish them accurately. Wrong amino acids can then be attached to tRNAs and be incorporated into proteins, creating mutant proteins. For this reason, some of the aaRSs have an important editing activity which can be encoded by a separate active site. This site clears the wrong amino acids attached to tRNAs (Nangle, Lagard et al. 2002; Schimmel 2008).

Baldwin and Berg (Baldwin and Berg 1966) stimulated the breakdown of misactivated Val-AMP on Isoleucyl-tRNA synthetase (IleRS) by adding Isoleucyl tRNA ($tRNA^{Ile}$) to cell extracts, indicating an intrinsic proofreading activity (later called editing activity) that could prevent misacylation of $tRNA^{Ile}$ with valine. Two ways of editing activity can occur; pre-transfer editing which occurs after an amino acid has been

activated, but before it has been attached to a tRNA. This involves hydrolysis of the misactivated aminocyl-adenylate, aa-AMP, and post-transfer editing which occurs after attachment of an amino acid to a tRNA and involves hydrolysis of aminoacyl-tRNA (aa-tRNA) (Ling, Reynolds et al. 2009; Martinis and Boniecki 2009).

Editing decreases the frequency of errors in the translation process and is important for translation quality control (Ling, Reynolds et al. 2009). When this activity is compromised, a wrong amino acid can be incorporated into proteins (Schimmel 2008). Studies in bacteria, human cell cultures and mice suggest that defects in editing are toxic to bacteria and can be associated to serious pathologies in human cells and mice (Ling, Reynolds et al. 2009).

The aaRSs are central to the translation process, but can also form a variety of multiprotein complexes which are present in all three domains of life whose functions are not yet fully elucidated (Figure 9). The aaRSs can also have noncanonical functions. For example, in humans GlnRS has an antiapoptotic function and the lysyl-tRNA synthetase (LysRS) is an inflammatory cytokine (Hausmann and Ibba 2008).

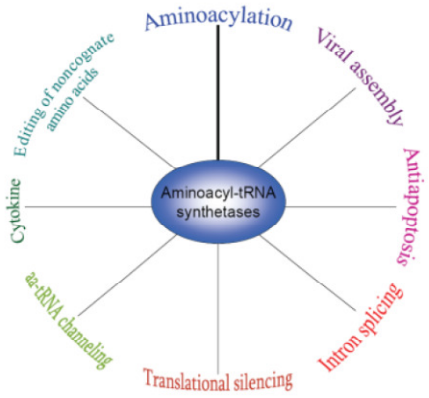


Figure 9. Functions that aminoacyl-tRNA synthetases can acquire in the three domains of life (Hausmann and Ibba 2008)

Due to the functional versatility of these enzymes, they can be associated with various human diseases and be attractive targets for the development of therapeutics. This will be topic of discussion in the subchapter Mistranslation and Disease.

1.3 Mistranslation

In spite of being a highly accurate process, protein synthesis is not error free. Aminoacylation errors are mainly caused by failure of the aaRS to differentiate between amino acids with similar chemical properties or by the incorrect recognition of tRNAs. Such errors are minimized by the aaRS editing mechanisms and by highly specific tRNA-aaRS interactions. At the ribosome level four main types of errors can occur, namely missense errors, which result in the substitution of one amino acid for another; nonsense errors, that cause readthrough of stop codons and produce proteins with extended C-termini; frameshifting errors, that alter the mRNA reading frame, producing out-of-frame truncated proteins; and finally processivity errors, which cause premature termination (Farabaugh and Bjork 1999; Moura, Carreto et al. 2009).

The frequency of global translational error *in vivo*, ranging from *E. coli* to mammalian cells, is in the order of 10^{-4} (Gallant and Palmer 1979; Moura, Carreto et al. 2009; Reynolds, Lazazzera et al. 2010). This physiological frequency of error is mitigated by protein quality control systems which destroy mistranslated proteins. Environmental conditions (for example amino acid starvation) can influence the fidelity of translation and the frequency of translation errors can increase significantly (Gallant and Palmer 1979; Moura, Carreto et al. 2009). But, mistranslation occurs at every step of protein synthesis, leading to synthesis of aberrant proteins (Reynolds, Lazazzera et al. 2010).

As already described, the mechanisms of quality control are responsible for maintaining the level of error low, but these mechanisms can be overloaded which may have important consequences for the cell. For example in mice an editing defective alanyl-tRNA synthetase (AlaRS) is responsible for neurodegeneration (loss of Purkinge cells) and ataxia (Lee, Beebe et al. 2006).

Surprisingly, certain errors can be tolerated and may even be advantageous. *Candida spp.* contains a $tRNA_{CAG}^{Ser}$ that can be aminoacylated with serine and leucine leading to ambiguous decoding of CUG codons. *Candida albicans* can tolerate up to 28% misincorporation of leucine without compromising the growth rate and in certain environmental conditions this increased level of mistranslation in *Candida albicans* can be advantageous (Santos, Cheesman et al. 1999; Gomes, Miranda et al. 2007). Indeed, ambiguous cells are able to survive in toxic environments containing cadmium, arsenate and hydrogen peroxide (Santos, Cheesman et al. 1999; Miranda, Silva et al. 2006). In

mammalian cells mistranslation can also be beneficial. Exposure of virus induces high level of mistranslation in mammalian cells due to the increased misincorporation of methionine at various non-cognate codons. Since methionine residues can protect proteins from ROS damage this type of mistranslation protects cells against oxidative stress (Levine, Mosoni et al. 1996; Reynolds, Lazazzera et al. 2010).

1.4 Protein Misfolding and Protein Quality Control

After biosynthesis by the ribosome, proteins must be converted into tightly folded compact structures in order to function. In cells there are large numbers of auxiliary factors that assist in the folding process, as folding catalysts (enzymes that catalyze the correct disulfide bond combination) and molecular chaperones (ensure a high degree of folding fidelity). The term "folding" refers to the processes by which the newly synthesized protein molecule folds into its unique three-dimensional structure (Dobson 2004; Ecroyd and Carver 2008; Herczenik and Gebbink 2008).

Sometimes proteins cannot fold properly, and become unfolded or misfolded. Misfolding can be due to amino acid misincorporation resulting from genetic mutations or errors in transcription, mRNA processing or translation and it also depends on environmental conditions such as increased temperature, high or low pH or oxidative agents. Because of the lack of three dimensional arrangement unfolded proteins are nonfunctional (Johnston, Ward et al. 1998; Herczenik and Gebbink 2008; Powers and Balch 2008). These proteins are destabilized and expose natively hydrophobic residues that seek nonpolar surface area in others destabilized proteins, causing protein-protein aggregation (Drummond and Wilke 2008). These aggregates can be unordered amorphous aggregates or highly ordered fibrils that are called amyloid fibrils (Johnston, Ward et al. 1998; Herczenik and Gebbink 2008).

Misfolded proteins may disrupt cellular membranes and with it ionic balances required for viability, by inserting themselves into the phospholipid bilayer. Also, misfolded proteins can overload the proteasome, molecular chaperones, and the endoplasmatic reticulum (ER), generate ROS, activate the unfolded protein response (UPR) and finally lead to cell death through apoptosis (Figure 10). In addition to these toxic effects misfolded proteins can also trigger inflammatory responses, as in the case of Alzheimer's disease where neuroinflamation occurs due to local stimulation of complement system (Nangle, Motta et al. 2006; Herczenik and Gebbink 2008).

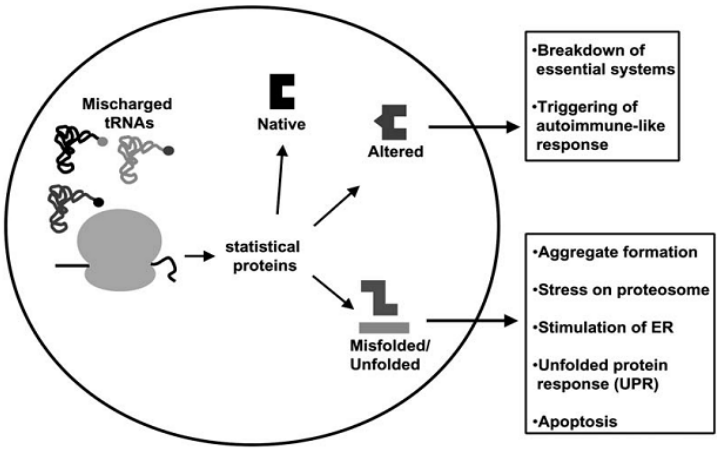


Figure 10. Possible consequences of global mistranslation. In this case mistranslation is caused by mischarged tRNAs (Nangle, Motta et al. 2006).

Protein quality control mechanisms are essential to cells and have evolved to supervise folding, counteract aggregation and eliminate misfolded and damaged polypeptide chains before they can exert toxic effects. These mechanisms comprise molecular chaperones, specialized intracellular proteases and accessory factors that regulate the activity of chaperones and proteases or provide communication between the various components (Gregersen, Bross et al. 2006).

After synthesis, proteins interact with chaperones to achieve their three dimensional conformation. There are many chaperones that have different functions: folding chaperones promote folding; holding chaperones lack ATPase activity and maintain solubility and finally unfolding chaperones contain two domains that harbor an ATPase and unfold misfolded proteins (Gregersen, Bross et al. 2006). Chaperones recognize selectively misfolded proteins and are the first step towards their elimination. Misfolded proteins can be refolded to a functional native state or can be sequestered by chaperones to prevent toxic interactions. Some chaperones such as Hsp70 and its cofactor Hsp40, seem to alleviate toxicity by sequestering soluble misfolded proteins or by modulating their conformation. In Huntington's disease these chaperones induce conformational rearrangements in Hungtintin, a pathogenic protein, and disfavor the accumulation of specific soluble polyQ fibril intermediates (Wacker, Zareie et al. 2004 ; Muchowski and Wacker 2005).

Proteins that cannot be refolded must be eliminated by proteases. The most prominent protease of the cytosol is the proteasome, but there are other proteases that belong to the protein quality control system, for example in mitochondria (Gregersen, Bross et al. 2006).

The ubiquitin-proteasome system (UPS) consists of a large number of proteins that identify, target and destruct proteins. This system can hydrolyze almost any polypeptide chain because it contains three different proteases with complementary activities: chymotrypsin-like, trypsin-like and caspase-like activities (Voges, Zwickl et al. 1999). These proteases are embedded in a large barrel-shape proteolytic complex, the proteasome core 20S. This proteolytic core can only be accessed through gated entrances on either side of the proteasome (19S subunits). This compartmentalized structure safeguards the highly specific nature of the process and prevents non-specific protein degradation (Dantuma and Lindsten 2010).

Proteins need to be marked for degradation by the proteasome. This is achieved by the covalent linkage of a chain of ubiquitin proteins. Specific enzymes, ubiquitin activase (E1), ubiquitin conjugases (E2s) and ubiquitin ligases (E3s) are responsible for the linkage of the C terminus of the 76 amino acid long ubiquitin polypeptide to the α-NH_2 group of an internal lysine residue of the substrate. Polyubiquitinated proteins bind to specific ubiquitin binding domains (UBD) which are located in the proteasome. Prior to degradation, in the proteasome, chains of ubiquitin are disassembled to ubiquitin monomers and the protein is unfolded and funneled through the narrow entrance of the proteasome to be hydrolysed (Figure 11) (Dantuma and Lindsten 2010).

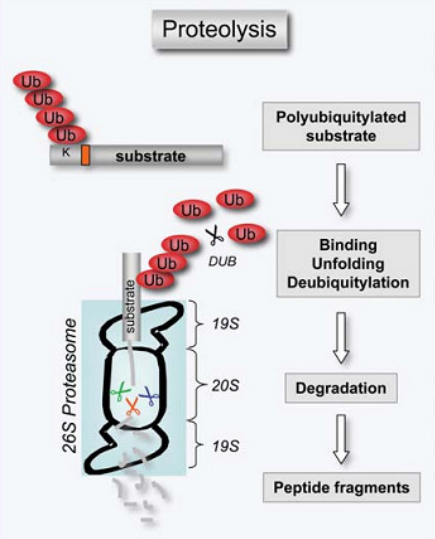


Figure 11. Protein degradation by the ubiquitin-proteasome system. Ub-ubiquitin; DUB-deubiquitylation enzymes(Dantuma and Lindsten 2010).

Misfolding proteins residing in the ER can also be degraded by proteasome, a process called ER-associated degradation (ERAD). Proteins in the ER are retrotranslocated to the cytosol, to be degraded by the UPS (Dantuma and Lindsten 2010).

If there is an increase in accumulation of misfolded proteins in ER and UPS is impaired, cells alleviate stress through the UPR. UPR consists in the transcriptional activation of genes required for protein folding, ER expansion ER-Golgi trafficking and ERAD, which all act collectively to relieve stress within the ER and reestablish its normal function. Translation of mRNAs is also initially inhibited, reducing the influx of new proteins into the ER (Haynes, Titus et al. 2004; Xu, Bailly-Maitre et al. 2005). Activation of UPR can be beneficial to cells, but if it is too prolonged UPR can lead to accumulation of ROS via the UPR-regulated oxidative protein folding machinery in the ER or in the mitochondria. In this way UPR contributes to cell death, usually apoptosis (Haynes, Titus et al. 2004).

In spite of not being normally considered a mechanism of protein quality control, cells can also degrade proteins through autophagy. Autophagy is the generic name used for any intracellular process that results in the degradation of cytosolic components inside the lysosomes. In these structures there is a complete and irreversible dissociation of the substrate into its essential constituents (Cuervo 2004).

Autophagy has been linked to some pathologies such as Alzheimer's disease, Parkinson's disease and different forms of cancer. This process is responsible for elimination of altered cytosolic constituents such as aggregated proteins and damage organelles and preserves cells from further damage (Larsen and Sulzer 2002). In this way, activation of autophagy could play a protective role in early stages of several diseases. Once a certain level of intracellular damage is reached, autophagy might instead become an efficient way of removing the injured cell from a tissue (Larsen and Sulzer 2002).

Therefore, cells can control the production quality proteins that constitute the proteome by diverse mechanisms. The first lines of protein quality control are the molecular chaperones that can refold proteins in the cytosol, ER or mithocondria. Then misfolded proteins in the cytosol or ER can be degraded by the ubiquitin-proteasome system (Gregersen, Bross et al. 2006). When ER stress occurs UPR is activated to counteract accumulation of misfolded proteins and aggregates. Finally misfolded proteins and aggregates can also be degraded by autophagy.

1.5 Diseases associated with mistranslation

Diseases have been associated with mistranslation. A clear case is the mistranslation associated to mutations in the editing domain of the mouse AlaRS which allow the enzyme to charge $tRNA^{Ala}$ with serine or glycine. Indeed, the editing is essential to deacylate Ser-$tRNA^{Ala}$ and Gly-$tRNA^{Ala}$. If these mischarged tRNAs are not cleared, serine or glycine will be incorporated instead of alanine, leading to the synthesis of misfolded proteins that are extremely toxic to neurons. Mice that have mutant AlaRS loose rapidly Purkinje cells in the cerebellum and develop severe ataxia (Park, Schimmel et al. 2008).

Mutations in genes encoding aaRSs can also lead to neuropathies as Charcot-Marie-Tooth (CMT) disease, which is characterized by muscular weakness and atrophy of the distal extremities. Genes that encode GlyRS and TyrRS have mutations in CMT patients that lead to conformational changes in the structure of these enzymes that prevent efficient interaction with the respective tRNAs. At least 10 mutant alleles have been associated to defective GlyRS (Xie, Nangle et al. 2007 ; Park, Schimmel et al. 2008).

Also mutations in tRNAs, specially in mitochondrial tRNA genes, are correlated with severe diseases, including fatal cardiopathies, encephalopaties, myophaties, diabetes and others. Mioclonic epilepsy with ragged-red fibers (MERRF) and mitochondrial myopathy, encephalopathy, lactic acidosis and stroke (MELAS) are the most common syndromes correlated with point mutations in mitochondrial tRNA genes. Eighty percent of patients with MELAS carry a maternally inherited A>G mutation in the nucleotide 3243 in the $tRNA^{Leu(UUR)}$ that affects its stability, impairs charging efficiency with Leu and prevents taurine modification in the anticodon. An A8344G mutation in the $tRNA^{Lys}$ causes MERF. This mutation affects aminoacylation and taurine modification of the wobble nucleotide of the anticodon which is essential for codon-anticodon pairing on the ribosome (Florentz and Sissler 2001; Rabilloud, Strub et al. 2002 ; Levinger, Morl et al. 2004; Scheper, Knaap et al. 2007).

Ribosome frameshifting can have a potential role in generating aberrant proteins such as ubiquitin B (UBB^{+}) and β-amyloid precursor protein (APP^{+}) which are implicated in Alzheimer's and other neurodegenerative diseases. These proteins are called "+1 proteins" because they have carboxyl-terminal amino acids encoded by an alternative reading frame of the mRNA. Aberrant proteins inhibit the function of the proteasome,

leading to their accumulation and can form neurofibrillary tangles and neuritic plaques (Wills and Atkins 2006) .

1.6 Cellular stress response due to misincorporation of amino acid analogues

There are hundreds of amino acid analogues produced by plants, synthesized *in vitro* or produced *in vivo* by oxidation of amino acid side-chains (Rodgers and Shiozawa 2008).

Certain amino acid analogues can have a similar structure to the natural amino acids and can escape detection by the cellular protein synthesis machinery, becoming misincorporated into the growing polypeptide chain of proteins. This incorporation leads to altered function of the proteins, production of misfolded or unfolded proteins and aggregates (Rodgers and Shiozawa 2008).

The aaRSs can select the correct amino acids, but if an amino acid analogue has a similar structure to the natural amino acid it may bind to an aaRS be activated and incorporated into proteins (Figure 12) (Richmond 1962).

Amino acid analogues misincorporation have been studied in bacteria and insects (plant amino acids show anti-microbial and insecticidal effects) but have been poorly studied in mammalian systems (Hendrickson, Crécy-Lagard et al. 2004; Rodgers and Shiozawa 2008). The response of cells to amino acid analogues depends on several factors, namely the efficiency of their transport into the cell; the extent to which the analogue satisfies the substrate specificity of the enzyme; the availability of the natural amino acid and the extent to which the incorporation of the analogue modifies the structure of the proteins and the enzymatic activity (Richmond 1962; Rodgers and Shiozawa 2008).

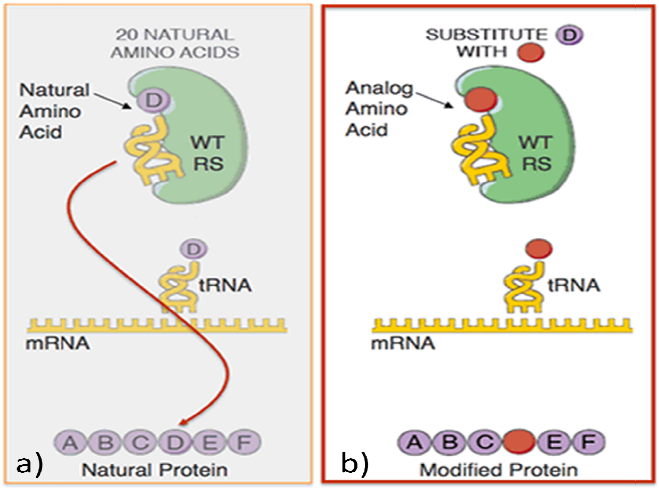


Figure 12. a) Incorporation of a natural amino acid into protein. b) Incorporation of an amino acid analogue into protein. The amino acid analogue will be recognized by aaRS and activated. Then the amino acid analogue will be transferred to the 3'end of tRNA and inserted into the protein in place of the natural amino acid, leading to the formation of a modified protein. WTRS-wild type aminoacyl-tRNA synthetase [(adapted from (ALLOZYNE 2010)].

The larger amino acids have the most effective analogues which suggest that the biological sites that deal with small amino acids, like glycine, serine and alanine are more stringent to molecular shape and size than those handling larger amino acids. Little changes in the structure of small amino acids can lead to a relatively larger change in overall shape and size than in a large amino acid (Richmond 1962). The amino acids and the analogues compete for the active site of the aaRS and in cases where there is no natural amino acid, preferential incorporation of the analogue will occur leading to severe toxic effects in cells (Rodgers and Shiozawa 2008).

Canavanine is a naturally occurring arginine analogue and is an effective substrate for arginyl-tRNA synthetase (ArgRS), which activates and aminoacylates it on the arginyl-tRNA ($tRNA^{Arg}$). The terminal methylene linked to the guanidine group of Arg is replaced by oxygen, in canavanine (Figure 13). The oxyguanidino group of L-canavanine (pK_a of 7,05) is less basic than the guanidine group of L-arginine (pK_a of 12,48), so proteins with L-canavanine incorporated, lack the capacity to form crucial ionic interactions that lead to disruption of tertiary and/or quaternary protein structure (Igloi and Schiefermayr 2009). The result is an altered protein function or disruption of enzymatic activity and a potentially rapid degradation of the proteins (Pines and Lindon 2005). By altering specific proteins as histones and heat shock proteins, cells exposed to canavanine have altered

DNA replication and translation mechanisms and decreased tolerance to heat, radiation and other stressors (Worthen, Chien et al. 1998).

Azetidine-2-carboxylic acid (AZC) can substitute proline (Pro) during protein synthesis. It leaks one carbon atom in its ring relative to Pro, which causes reduction of thermal stability of proteins and misfolding (Figure 13). Collagen contains large amounts of Pro, so is particularly sensitive to the replacement of Pro with AZC (Rodgers and Shiozawa 2008). The replacement of Pro with AZC causes growth inhibition of *E. coli* and in yeast it induces protein misfolding and selectively activates the heat shock factor which is required for the subsequent G_1 arrest of the cell cycle (Trotter, Kao et al. 2002 ; Jhon and Kang 2007).

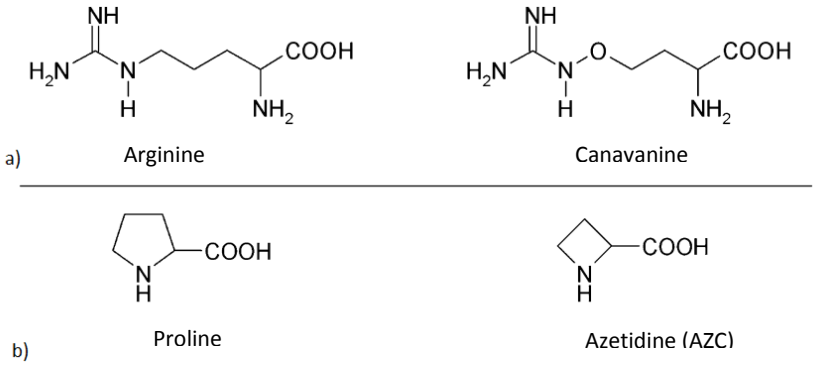


Figure 13. Structure of natural amino acids and respective analogues. a) Natural amino acid Arginine and its analogue Canavanine. b) Natural amino acid Proline and its analogue Azetidine (Rodgers and Shiozawa 2008).

Canavanine only differs from arginine in the structure of its side chain, so it is not expected to significantly alter the conformation of the polypeptide backbone into which it is incorporated. Therefore treatments with canavanine cause less protein misfolding than those with AZC (Trotter, Kao et al. 2002).

The incorporation of amino acid analogues originates modified proteins that tend to degrade faster than native proteins in bacterial and human cells and have been shown to induce many heat shock proteins (HSPs) (Rodgers and Shiozawa 2008).

The use of amino acid analogues could therefore be useful to understand the response of cells to mistranslation, which mechanisms are required to protect cells against the accumulation of damaged proteins and how these damaged proteins can lead to cellular

death. Some of these mechanisms have already been identified for yeast but are less well studied in human cells (Figure 14).

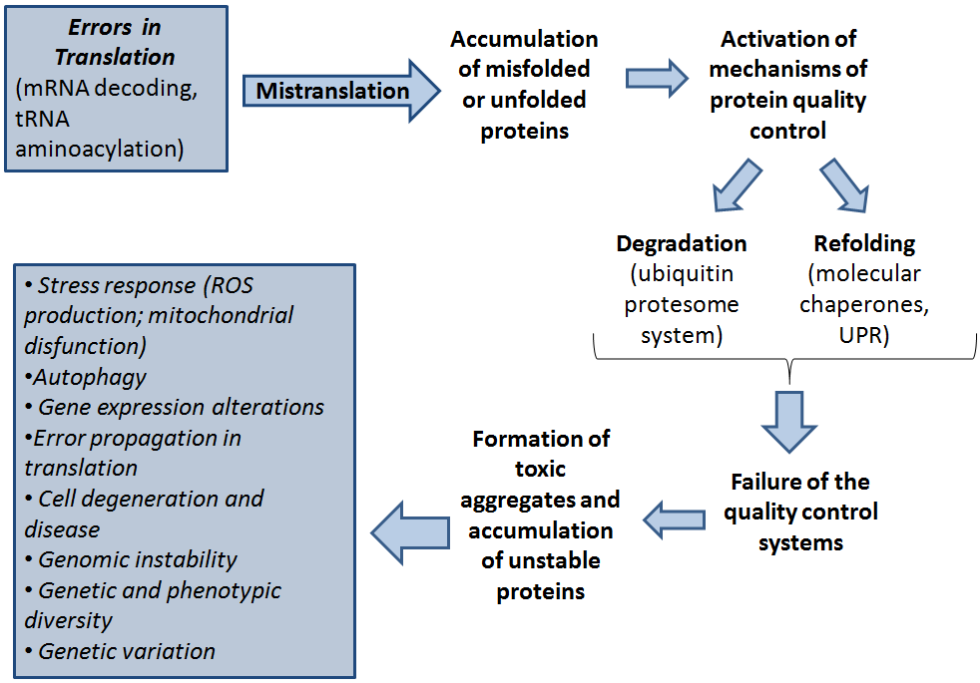


Figure 14. Incorporation of amino acid analogues can lead to synthesis of aberrant proteins that misfold or unfold and activate mechanisms of protein quality control. For example: molecular chaperones, ubiquitin-proteasome system or UPR-unfolded protein response, among others. The failures of these mechanisms lead to formation of toxic aggregates and unstable proteins which originate many responses by the cell to the mistranslation phenomenon.

1.7 Aims of the study

The aim of this Master thesis was to study the cellular responses to mistranslation, using the human cell line HEK293FT. These cells were exposed to amino acid analogues, Azetidine-2-carboxylic acid (AZC) and canavanine, which are known to misincorporate into proteins during protein synthesis. This study is important to understand the response of human cells to mistranslation and to start unveiling the cellular and molecular mechanisms of diseases associated to protein misfolding and aggregation.

The biological questions addressed here are the following:

1- Which molecular mechanisms are activated in response to mistranslation?

2- How is the transcriptional response to mistranslation regulated?

3- Are HEK293FT cells a good model system to study cell degeneration and disease associated to mistranslation?

Chapter II – Materials and Methods

2.1 Materials

HEK293FT cells were purchased from Invitrogen (Barcelona, Spain). Dulbecco's modified Eagle's medium (DMEM) as well as poly-D-lysine, bovine serum albumin (BSA), paraformaldehyde, L-canavanine and L-azetidine-2-carboxylic acid (AZC) were purchased from Sigma-Aldrich (Sintra-Portugal). Fetal bovine serum (FBS), glutamine, trypsin and geneticin (G418) were purchased from Gibco, Invitrogen Life Technologies (Barcelona, Spain). The BCA assay kit was purchased from Pierce, as a part of Thermo Fisher Scientific (Rockford, Illinois, USA). Propidium Iodide was purchase from Cytognos, as a part of the kit Cycloscope B-All for DNA studies. Hoechst 33342 trihydrochloride trihydrate was purchased from Molecular Probes (Leiden, Netherlands). The reagents to perform Oligo Microarrays were purchased from Agilent Technologies (Santa Clara, USA). Anti-ubiquitin antibodies were purchased from Dako (Aachen, Germany), and normal rabbit IgGs were purchased from Santa Cruz Biotechnology (California, USA).

All other reagents were from Sigma-Aldrich (Sintra-Portugal) or from Merck (Darmstadt, Germany). L-canavanine and L-azetidine-2-carboxylic acid (AZC) were kept in aqueous stocks.

2.2 Methods

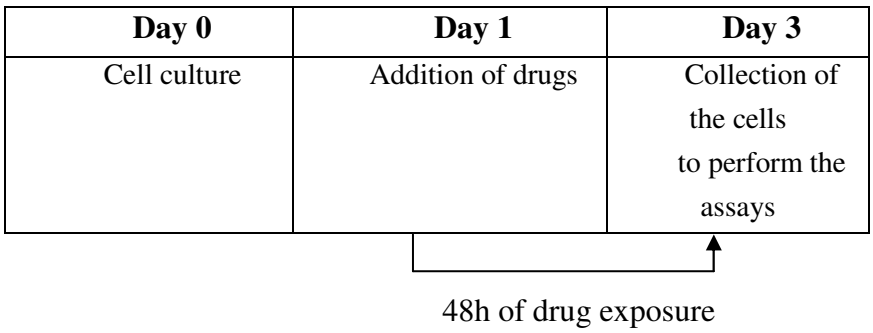
2.2.1 Cell Culture

Human embryonic kidney (HEK) 293FT cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) with 10% of fetal bovine serum (FBS), 5% geneticin, 1% L-glutamine and supplemented with non-essential amino acids (10mM). Cells were kept in falcon flasks (75cm^2) at 37ºC, in a humidified incubator with 5% CO_2/95% air.

2.2.2 Preparation of cells and stimulation with AZC and canavanine

Single cell suspensions were obtain by trypsinization (1mL, 5 min at 37ºC) of monolayer cultures. Then 5mL of DMEM were added to stop trypsin (0,25%) activity and cell suspensions were transferred to a falcon tube. Cell density in the suspensions was determined by counting the number of cells using a Neubauer chamber.

Cells were cultured at the density of 0.3x10^{5}cells/cm^2 for all the assays, and were maintained for 24h before incubation with the drugs for 48h.



2.2.3 Cellular viability assay – MTT

To assess cell viability the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) reduction test was used. When taken up by living cells, MTT is converted from yellow to a water-insoluble blue-colored precipitate, formazan, by mitochondrial succinate dehydrogenase and probably by pyridine nucleotide cofactors NADH and NADPH in the cytoplasm of metabolically active cells (Wyllie 2009).

MTT(1mg/mL) – in sodium buffer with 1mM $CaCl_2$, was added to the cultures after 48h of drug exposure, and incubated for 2h30min at 37ºC in humidified incubator with 5% CO_2/95% air. Formazan crystals were dissolved in 0.04M HCl in isopropanol. Solutions were transferred to a 96 multiwell plate to read absorbance in an ELISA plate reader Spectra Max plus (Molecular Devices). A plate analyzer with dual wavelength measuring system was used, at a test wavelength of 570nm and a reference wavelength of 620nm. All experiments were carried out in triplicate, and the reduction of MTT was expressed in the experimental conditions as percentage of that observed in the control cultures (100%).

2.2.4 Fluorescence Microscopy

HEK293FT cells were plated on coverslips (12mm) coated with poly-D-lysine (0,1mg/ml). After incubation in the presence or in the absence of the drugs, the cells were fixed with 4% paraformaldehyde/4% sucrose for 15 min at room temperature. The cells were then washed 3x with phosphate buffered saline (PBS) and stained with the fluorescent dye Hoescht 33342 (1 μg/ml in sodium buffer) for 15 min at room temperature. The coverslips were mounted on glass slides with Dako mounting medium and examined with a Zeiss Axiovert 200 fluorescence microscope (20× objective).

Hoechst is a permeable DNA stain that binds preferentially to A-T base-pairs. This dye is excited by ultraviolet light at around 350nm and emits blue fluorescence. By microscopic observation, nuclei of viable and non viable cells can be distinguished.

Nuclei of viable cells and non viable cells were counted to calculate the percentage at each microscopic field. All experiments were carried out in triplicate and the results were expressed in percentage of viable or non viable cells for each experimental condition.

2.2.5 Flow Cytometry

Flow cytometry is a rapid technique that allows characterizing different cells properties by analyzing individual cells in a population. Two of the applications of this technique are the measurement of cellular DNA content and the analysis of the cell cycle.

For these analyses, cells were stained with a fluorescente dye, Propidium Iodide (PI) that intercalates between the bases of the DNA with no sequence preference.

Cells were cultured in a 6 multiwell plate coated with poly-D-lysine. After 48h of drug exposure, cell suspensions were obtained by trypsinization of the cells (1mL, 5 min at 37ºC) and the enzyme was inactivated by 500µL of 1% PBS and 100µL of 10% FBS. Cell suspensions were centrifuged (0.2g, 5min at 4ºC) and the supernatant was discarded. The pellet was resuspended in 1mL of 1% PBS. The centrifugation step was repeated and the pellet was resuspended in 100µL of 1% PBS. Cells were kept at 4ºC until data acquisition by the flow cytometer.

100µL of sample were transferred to cytometer tubes and samples were gently aspirated several times with a Pasteur pipette to obtain a cell suspension with minimal cell aggregation. 200µL of PI was added to each sample tube. Tubes were incubated horizontally in the dark for 15min. Before acquisition each sample was resuspended in 500µL of 1% PBS. About 20.000 events were acquired in the cytometer FACS Calibur (Becton-Dickinson) using Cell Quest software (PMac) and analyzed using ModFitLT^{TM} v2.0 software (PMac). The results were expressed in percentage of cells in each phase of the cell cycle (G0-G1; G2-M; S) for each experimental condition.

2.2.6 Proteasome activity assay

Proteasome is responsible for the degradation of proteins that are damaged or not necessary for the cell, by proteolysis. The 26S proteasome is an ATP-dependent proteolytic complex which contains a core particle (20S) and two regulatory caps (each one 19S). In the core, enzymes with specialized activities (chymotripsin-like, trypsin-like and caspase-like) break down peptide bonds (Dantuma and Lindsten 2010). Proteasome activity can be measured using a labeled peptide (Suc-LLVY-AMC) which is cleaved by enzymes with chymotrypsin-like activity thereby generating a strong fluorescent signal.

Cell extracts were prepared to detect proteasome activity in experimental conditions. Cells were washed with ice-cold PBS. A lysis buffer (Appendix 1) was added to the cells, which were then scraped. The extracts were sonicated and left on ice for 5 min. The lysates were then centrifuged at 12000xg for 10 minutes at 4ºC and the supernatants recovered. Protein was quantified using the Bio-RAD method, and the extracts were diluted in lysis buffer (Appendix 1) to equalize protein concentrations of all samples. 20µg

of protein were incubated in a 96 multiwell plate with proteasome activity buffer (in duplicate) (Appendix 1) in a final volume of 100µL.

The fluorescence intensity was monitored in a Spectra Max (Molecular Devices) fluorometer for one hour (measurements in every 5 min), using the wavelengths: 380nm excitation and 460nm emission.

We carried out the experiment four times and proteasome activity was expressed in the experimental conditions as the percentage of that observed in the control cultures (100%).

2.2.7 Western blot

In order to analyse protein ubiquitination, samples were subjected to sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) where protein migration is determined by their molecular weight. SDS is an anionic detergent that denatures proteins by wrapping around the polypeptide backbone and confers a negative charge to the polypeptide in proportion to its length.

The cells were washed with ice-cold PBS, before addition of lysis buffer (Appendix 1). The lysed cells were then scraped and the extracts were sonicated and left on ice for 30-40min. The lysates were then centrifuged at 12000xg for 20 minutes at 4ºC and the supernatants recovered. The protein concentration was quantified using the BCA protein assay kit, and then the extracts were diluted with a 5x concentrated denaturing buffer (Appendix 1) and denaturated at 100ºC for 10min.

Denatured extracts were resolved on SDS-PAGE (8-10% acrylamide) (Appendix1) and then electrotransferred onto a Polyvinylidene Difluoride (PVDF) membrane, overnight at 40V. 60µg of protein were loaded into the gels. Membranes were blocked for 1h in Tris-buffered saline with 0.1% Tween20 (TBS-T) and 5% non-fat dry milk (Appendix 1) and then incubated for 1h at room temperature, or overnight at 4°C, with the primary antibody (anti-ubiquitin 1:2500) in TBS-T containing 1% milk. Membranes were then washed and incubated with alkaline phosphatase conjugated IgG secondary antibody (according to the primary antibody host specificity; anti-rabbit 1:10000) in TBS-T with 0.5% milk for 1h. After additional washes, the membranes were developed using enhanced chemifluorescence (ECF) substrate, and scanned using a Storm 860 Gel and Blot Imaging System (Amersham Biosciences, Buckinghamshire, UK). The intensity of the bands on the

membrane was analyzed with ImageQuant 5.0 software. For multiple probing, membranes were stripped of antibodies with NaOH 0.2M for approximately 20min, at room temperature, with gentle shaking, and then blocked with TBS-T with 5% non-fat dry milk and reprobed with primary (anti-actin 1:5000) and secondary (anti-mouse 1:10000) antibodies.

2.2.8 RNA extraction

To perform microarrays, RNA extraction was carried out using Trizol reagent.

Cells were washed twice with PBS and then 1mL of Trizol was added to each well of 6 multiwell plates. Samples were incubated for 5 min at room temperature to permit the complete dissociation of nucleoprotein complexes. Cells were scrapped and the content of each well was transferred to an eppendorf RNase free tube. 200 µL of chloroform (200µL/mL Trizol) was added to each eppendorf which were vigorously shaken by hand for 15 seconds and incubated for 2 to 3 min at room temperature. Samples were centrifuged at 12000 g for 15 min at 4ºC. The supernatants were then transferred to new eppendorf tubes and 500 µL of isopropanol (0.5 mL isopropanol/mL Trizol) was added to each eppendorf. Samples were centrifuged at 12000 g for 10 min at 4ºC. The supernatants were discarded and the RNA pellets were washed with 1 mL of ethanol 75% (1 mL ethanol/mL Trizol). Samples were centrifuged 12000 g for 5 min at 4ºC. The supernatant was discarded and the pellet was resuspended in 10 µL of H_2O miliQ (DEPC). Eppendorfs were incubated at 60ºC for 10 min.

2.2.9 Oligonucleotide Microarrays

A microarray is a special array of oligonucleotide probes, closely arranged in a small solid support surface, the slide. Oligonucleotide probes are arranged in spots and each spot corresponds to a specific gene. The position and nucleotide sequence of each probe are known. These arrays allow the simultaneous examination of the expression of thousands of genes.

RNA sample preparation, hybridization, microarray wash, scaning and feature extraction were carried out using the protocol “One-Color Microarray-Based Gene

Expression Analysis – Low Input Quick Amp Labeling" version 6.5 from Agilent Technologies. In this protocol Cyanine 3-labeled targets were used to measure gene expression in experimental and control samples. No alterations were made to the original protocol.

Raw data were normalized using BRB- Array Tools (Simon, Korn et al. 2003a). One-way ANOVA (p-value 0.01) was performed using MeV software (Saeed, Bhagabati et al. 2006). M values (log2 Ratios) were calculated using control samples as a reference. Functional analysis of expression data was carried out using DAVID Bioinformatics Resources 6.7, an online database (Huang, Sherman et al. 2008; Huang, Sherman et al. 2009).

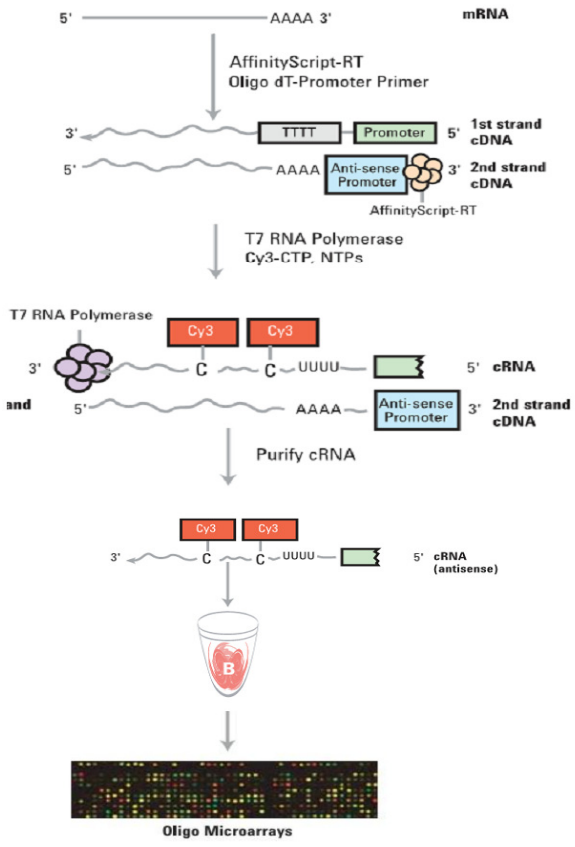


Figure 15. Schematic of amplified cRNA procedure (adapted from the protocol: "One-Color Microarray-Based Gene Expression Analysis – Low Input Quick Amp Labeling" version 6.5 from Agilent Technologies). From mRNA is produced cDNA using T7 promoter primer. From cDNA, cRNA attached to cyanine 3 (fluorescent dye) is produced. After purification cRNA will be hybridized in the oligo microarray.

2.2.10 Statistical analysis

Results are presented as means+SD of the number of experiments indicated. Statistical significance for the results of cellular viability assay (MTT) and proteasome activity assay was assessed by one-way ANOVA analysis followed by the Dunnett's test, using the software package GraphPad Prism 4. Statistical significance of the results of fluorescence microscopy was assessed by one-way ANOVA analysis followed by the Kruskal-Wallis test, using GraphPad Prism 4. Statistical significance of the results of flow cytometry was assessed by two-way ANOVA analysis followed by the Bonferroni test, also using GraphPad Prism 4.

Chapter III – Results

3.1 Effect of mistranslation on cellular viability and proliferation

3.1.1 Cellular viability and proliferation – MTT results

The MTT assay allows a rapid evaluation of the metabolic activity of viable cells and in toxicity studies using cell lines the changes in the reduction of MTT may arise from alterations in the rate of cell proliferation and/or effects on cell viability.

To first assess the effect of the amino acid analogues, AZC and canavanine on cell viability/proliferation, the MTT assay was performed after exposure of the cells to the amino acid analogues for 48h. The formazan product of this assay was quantified by spectrophotometry in an ELISA plate reader, as described in methods. The mean of absorbance values for each condition was calculated as well as the percentage of MTT reduction relative to the control (cells without addition of amino acid analogues). AZC and canavanine were tested at the following concentrations: 500µM; 1mM; 1.25mM; 1.5mM; 2mM and 2.5mM (Figure 16).

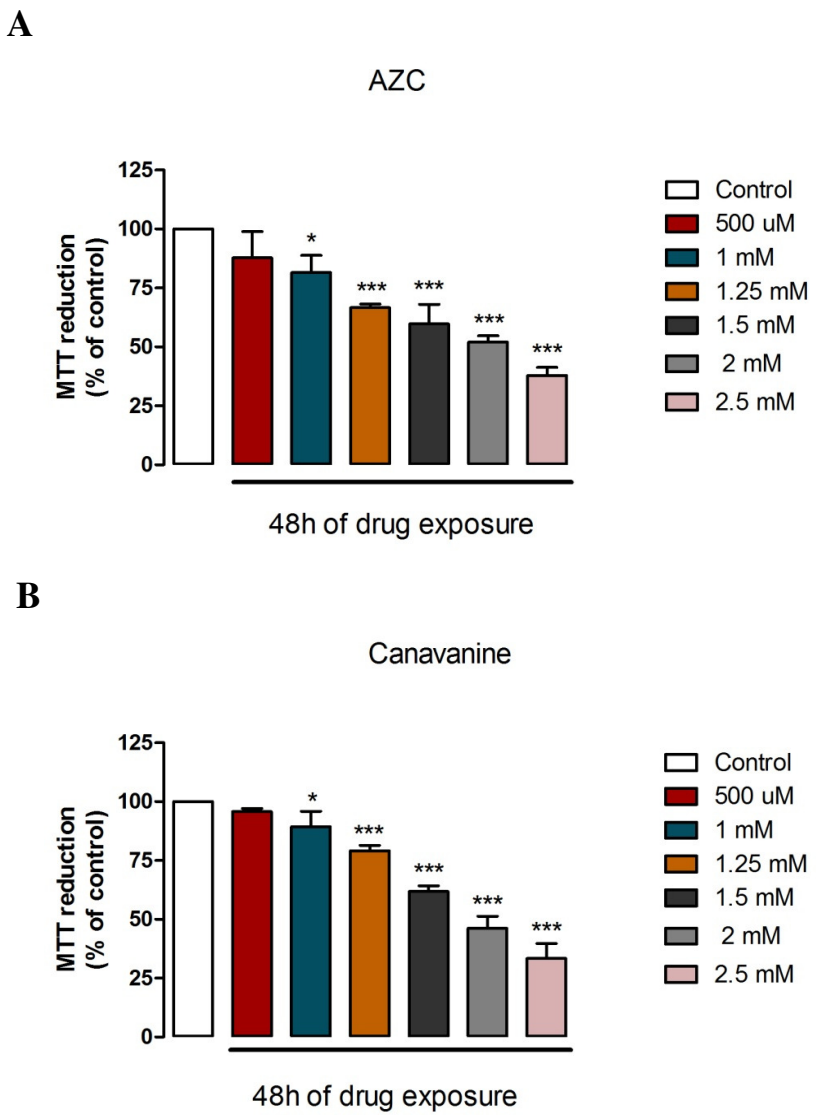


Figure 16. Effect of AZC (A) and canavanine (B) on cellular viability/proliferation. Percentage of MTT reduction for each condition is relative to the control. Values represent mean+SD of triplicates. Statistical analyses were performed using one-way ANOVA and Dunnett´s post test (* represents $p<0.05$, *** represents $p<0.001$).

Both drugs decreased MTT reduction in a dose-dependent manner when compared to the control cells which were not exposed to the amino acid analogues.

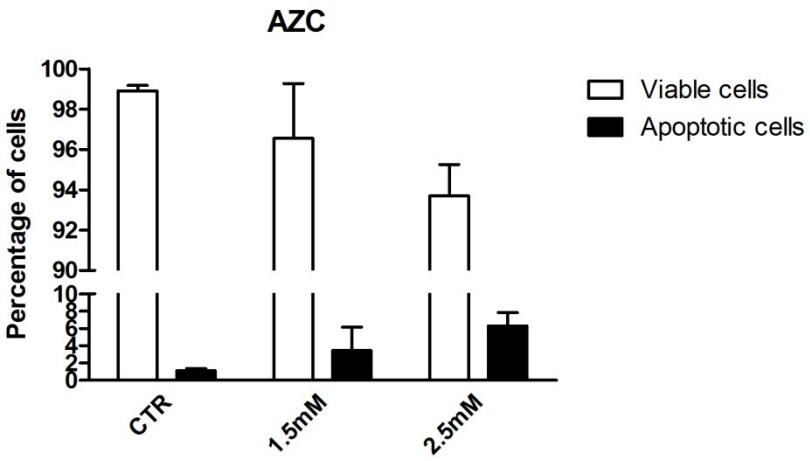
At a concentration of 500µM none of the amino acid analogues affected significantly the percentage of MTT reduction ($p<0.05$), however, when tested at higher concentrations, AZC and canavanine decreased significantly the percentage of MTT reduction (* $p<0.05$ and *** $p<0.001$). Results with AZC treatment were similar to those

of the canavanine treatment and may be attributed to a decrease in the rate of cell proliferation or a reduction in overall cell viability.

3.1.2 Cellular viability and proliferation assessed by nuclei counting and flow cytometry

Nuclei of viable and non-viable cells were stained with the fluorescent dye Hoechst 33342 and counted to distinguish if cells were dying, or if drugs were having an effect in cellular proliferation, instead of inducing cell death. Two concentrations of the amino acid analogues were tested, namely 1.5mM and 2.5mM. Cells were exposed to these concentrations for 48h except control cells which were not exposed to the amino acid analogues.

A



B

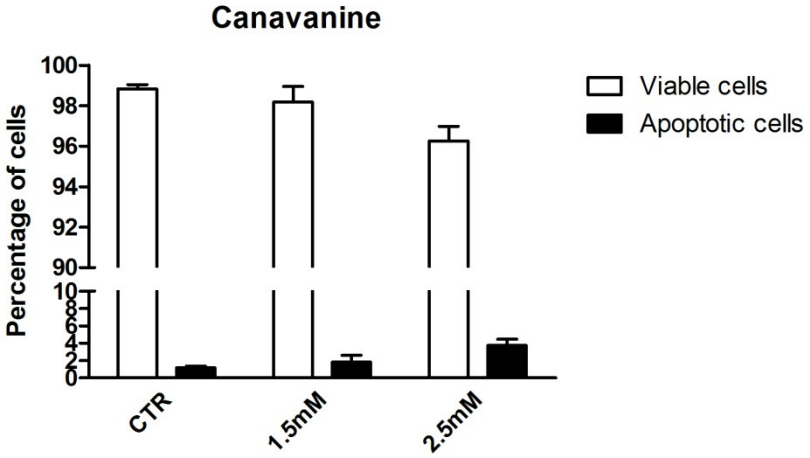


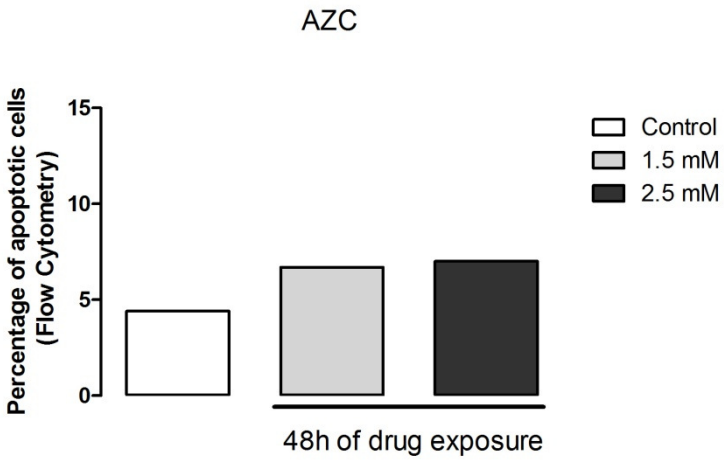
Figure 17. Effect of AZC (A) and canavanine (B) on cell viability. Nuclear morphology was analyzed by fluorescence microscopy after staining with Hoechst 33342 and the percentage of viable and apoptotic cells was then calculated. Values represent mean+SD of triplicates. Statistical analyses were performed using one-way ANOVA and Kruskal-Wallis post test (values were not statistically significant, $p>0.05$). CRT – control.

At the concentrations tested, AZC and canavanine showed very low toxicity. These amino acid analogues had a low effect inducing cell death by apoptosis (Figure 17). The percentage of viable cells decreased, with the increase in concentration of AZC (Control – 98,91%; 1.5mM – 96,56%; 2.5mM – 93,70%) and canavanine (Control – 98,84%; 1.5mM – 98,18%; 2.5mM – 96,26%). However, this difference was very small and not significant ($p<0.05$). In the same way, the percentage of apoptotic cells increased with AZC (Control

– 1,09%; 1.5mM – 3,44%; 2.5mM – 6,3%) and canavanine (Control – 1,16%; 1.5mM – 1,82%; 2.5mM – 3,74%), but not significantly ($p<0.05$).

Flow cytometry analysis with propidium iodide confirmed these results (Figure 18). Again, cells were exposed for 48h to 1.5mM and 2.5mM of the amino acid analogues. AZC and canavanine induced cell death but at a low level for the concentrations tested. The percentage of apoptotic cells assessed by flow cytometry were similar to the results obtained by nucleus cell counting. Percentage of apoptotic cells increased with increasing concentrations of the amino acid analogue (AZC Control – 4,4%; AZC 1.5mM – 6.69%; AZC 2.5mM – 7,01%; canavanine Control – 2,19%; canavanine 1.5mM – 6,63%; canavanine 2.5mM – 9,42%).

A



B

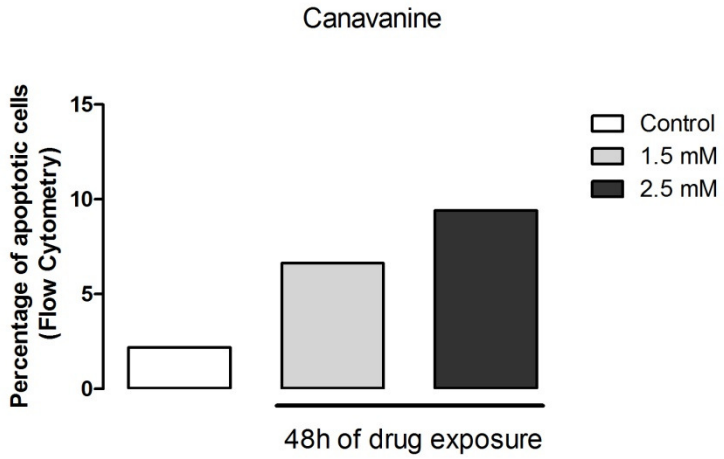
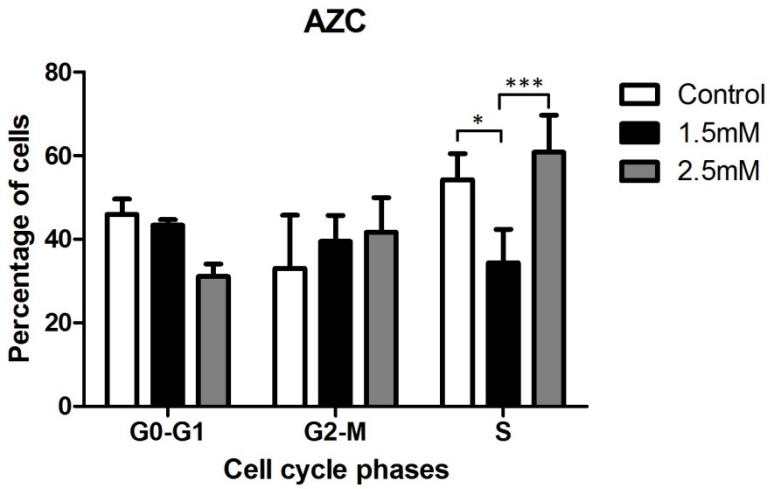


Figure 18. AZC (A) and canavanine (B) toxicity in HEK293FT cells, as determined by flow cytometry in nuclei stained with propidium iodide. Values represent one independent experiment, one replicate.

3.1.3 Cell cycle analysis by flow cytometry

In order to evaluate the effect of mistranslation on cell cycle progression, propidium iodide was also used to stain cells for further flow cytometry analysis. HEK293FT cells were also incubated for 48h with AZC or canavanine at 1.5mM and 2.5mM (Figure 19).

A



B

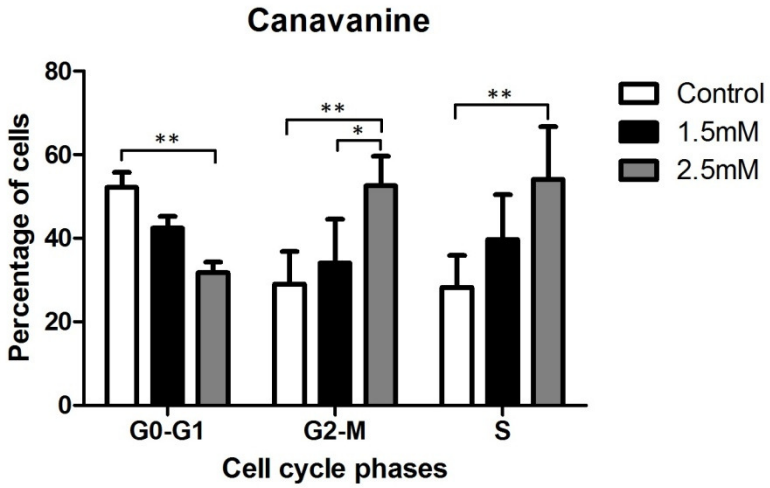


Figure 19. Effect of AZC (A) and canavanine (B) on HEK293FT cell cycle progression. Cell cycle analysis was conducted by flow cytometry with propidium iodide. After data acquisition percentage of cells in each phase of the cell cycle was calculated using ModFitLT^{TM} v2.0 software. Values represent mean+SD of four independent experiments. Statistical analyses were performed using two-way ANOVA and Bonferroni post test (* represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$).

The percentage of cells in phase G0-G1 decreased with the increase in the concentration of amino acid analogues. This difference was more proeminent in cells exposed to canavanine than AZC, where it was statistically significant. In contrats the percentage of cells in phase G2-M increased with increasing concentration of AZC and canavanine being this result statistically significant for canavanine. This can indicate that

cells are being arrested in phase G2-M. The observed changes in the percentage of cells in phase S, were different in HEK293FT cells incubated with AZC or canavanine. AZC decreased the percentage of S-phase cells (1.5mM), but increased it in presence of 2.5mM of AZC. Increased concentrations of canavanine increased the percentage of S-phase cells (statistically significant at $p<0.01$).

3.2 Protein quality control mechanisms activated in response to accumulation of misfolded proteins

3.2.1 Detection of proteins conjugated with ubiquitin by western blot analysis

Misfolded proteins are targeted, by the addition of ubiquitin chains to degradation by the proteasome system. In order to determine whether mistranslation induced formation of ubiquitinated proteins western blot analysis were carried out. Protein extracts from control cells and cells exposed for 48h to 1.5mM and 2.5mM of amino acid analogues were resolved by SDS-PAGE and then electrotransferred onto a PVDF membrane. Ubiquitin conjugates were detected using an anti-ubiquitin antibody.

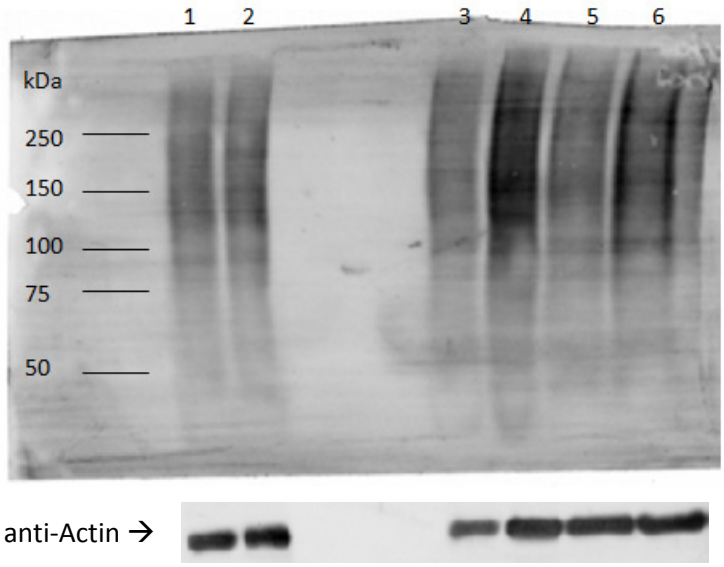


Figure 20. Effect of incubation of AZC and canavanine on protein ubiquitination. HEK293FT cells were treated with the drugs for 48h and anti-ubiquitin conjugates were detected with an anti-ubiquitin antibody. Samples 1 and 2 – Controls; 3 – AZC 1.5mM; 4 – AZC 2.5mM; 5 – canavanine 1.5mM; 6 – canavanine 2.5mM. SDS-PAGE (8-10% acrylamide).60µg of protein were loaded for each sample. Results of one independent experiment.

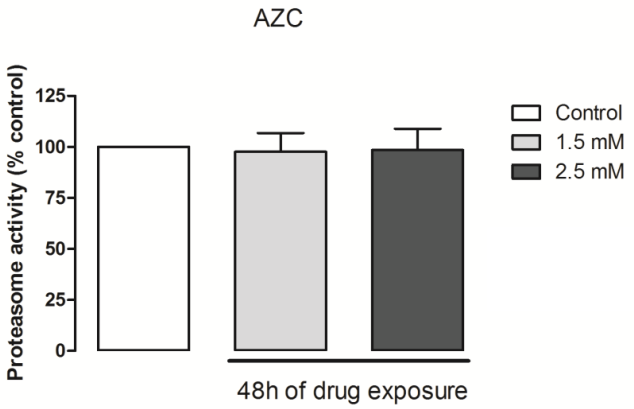
The results showed an increase in ubiquitinated proteins, particularly in the presence of 2.5mM AZC, where the antibody staining was stronger. Canavanine (2.5mM) also increased antibody staining of proteins, but to a less extent than AZC. Low

concentrations of AZC and canavanine (1.5mM) did not significantly changed the antibody staining relative to the control cells extracts (Figure 20).

3.2.2 Proteasome activity

In order to elucidate whether the amino acid analogues altered proteasome activity, the activity of the latter was determined using a proteasome activity assay. Concentrations of amino acid analogues and time of exposure were those used to detect the ubiquitin conjugates by western blot analysis (see above).

A



B

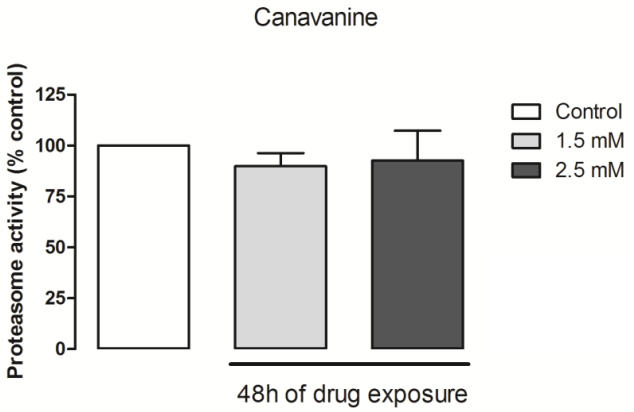


Figure 21. Effect of AZC (A) and canavanine (B) on proteasome activity in HEK293FT cells. Proteasome activity is expressed in the experimental conditions as a percentage of control (100%). Values represent mean+SD of four independent experiments. Statistical analyses were performed using one-way ANOVA and Dunnett´s post test (values were not significant, $p<0.05$).

Amino acid analogues at these concentrations and time of exposure did not affect significantly the activity of the proteasome (Figure 21), indicating that the accumulation of the ubiquitinated proteins is not due to a down-regulation of the proteolytic activity of the proteasome. AZC increased proteasome activity from 97,70% (1.5mM) to 98,52% (2.5mM), but this small increase was not statistically significant. The same happened for canavanine where proteasome activity increased slightly from 89,80% (1.5mM) to 92,65% (2.5mM), but again this increase was not statistically significant.

3.3 Transcriptional responses to mistranslation

In order to unveil the molecular meachanisms that are activated in response to mistranslation, oligo microarrays were performed. Once again we exposed HEK293FT cells to the amino acid analogues AZC and canavanine for 48h, using two concentrations 1.5mM and 2.5mM. Control cells were not exposed to the amino acid analogues. Total RNA was extracted from three independent cultures. After purification and labeling with Cy3, cRNA was hibridyzed on an Agilent Microarray as described in methods.

After extraction using Agilent Feature Extraction Software, the data was normalized using BRB Array Tools. In order to identify genes whose expression changed significantly in presence of the amino acid analogues, One-way ANOVA (p-value 0.01) was carried out using the MeV software package. Genes which expression was significantly different (p-value 0.01) were clustered in up-regulated (from control to 1.5mM and 2.5mM) and down-regulated genes (from control to 1.5mM and 2.5mM) for both amino acid analogues, using Expander 5 sofware. Genes that did not show the same pattern in both concentrations, for example were up-regulated from control to 1.5mM and down-regulated from 1.5mM to 2.5mM, in spite of having a significantly different expression pattern were not considered for the analysis.

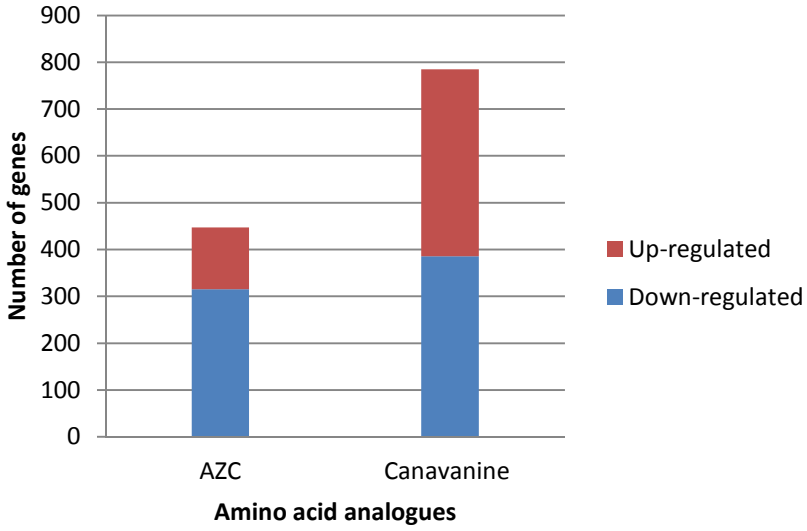


Figure 22. Number of genes up and down-regulated with the amino acid analogues AZC and canavanine.

Addition of AZC to cells altered expression of 447 genes, 132 of them were up-regulated and 315 were down-regulated, while canavanine deregulated 785 genes, 399 were up-regulated and 386 were down-regulated (Figure 22).

3.3.1 Functional analysis of down-regulated genes

Genes were grouped in: AZC up-regulated genes, AZC down-regulated genes, canavanine up-regulated genes and canavanine down-regulated genes. Functional analyses of each one of these groups was carried out using the online database DAVID Bioinformatics Resources 6.7.

Table 1. Functional categories of the genes down-regulated by AZC.

Gene Ontology Annotation	p-value (Modified Fisher Exact p-value, EASE score)
Extracellular matrix part	$1.1x10^{-3}$
Proteinaceous extracellular matrix	$1.3x10^{-3}$
Laminin complex	$3.4x10^{-3}$
Basal lamina	$1.2x10^{-2}$
Growth factor binding	$1.2x10^{-2}$
Nucleotidyltransferase activity	$2.2x10^{-2}$
Basement membrane	$4.3x10^{-2}$
Response to oxygen levels	$4.3x10^{-2}$
Collagen	$4.7x10^{-2}$
DNA polymerase activity	$5.0x10^{-2}$
Cell projection organization	$5.6x10^{-2}$
Oxidation reduction	$7.5x10^{-2}$
Cell morphogenesis involved in differentiation	$7.8x10^{-2}$

The incorporation of AZC into proteins decreased the transcription of genes encoding proteins of the extracellular matrix and mainly involved in cell-adhesion (Table 1). For example COLGA1 and COL13A1 codify two types of collagen, an important molecule in cell attachment, migration and organization of cells into tissues. The EGFLAM gene encodes a protein involved in matrix assembly and cell adhesiveness and is also in this group. Proteinaceous extracellular matrix proteins include proteins which are important for interaction between cells and between cells and extracellular matrix. An important gene in this process is the KAZALD1 which leads to matrix assembly. Genes that codify laminins, namely LAMA1, LAMA5, LAMC2, which are components of the basal lamina were also down-regulated. Another category of down-regulated genes is involved in response to oxygen levels. Surprisingly, some genes whose proteins are important to counteract the negative effects of ROS and promote inhibition of growth and cell death were down-regulated, namely DDIT4 and CDKN1A.

Some down-regulated genes were not annotated in any category, namely genes that code for several transmembrane proteins, and some members of the MAP kinase family.

Similar results were obtained with canavanine (Table 2). Genes involved in the maintenance of the extracellular matrix (Extracellular region part; Extracellular matrix structural constituent and Proteinaceous extracellular matrix; Extracellular matrix organization) were down-regulated, for example, genes that codify collagen (COL5A2, COL13A1, COL2OA1), laminin (LAMC2) and cell adhesion proteins (SPON2). Similarly to AZC, some genes that were down-regulated with canavanine are involved in the response to oxygen levels, namely the gene DDIT4, which is down-regulated in cells exposed to AZC. Genes that were not clustered in any GO term were involved in transcription, for example TCEA2 (transcription factor).

Table 2. Functional categories of the genes down-regulated by canavanine.

GO Annotation	p-value (Modified Fisher Exact p-value, EASE score)
Extracellular region part	$1.4x10^{-4}$
Extracellular matrix structural constituent	$1.8x10^{-3}$
Proteinaceous extracellular matrix	$3.0x10^{-3}$
Collagen	$6.8x10^{-3}$
Nucleotidyltransferase activity	$2.2x10^{-2}$
Glycoprotein metabolic process	$1.7x10^{-2}$
Nucleobase, nucleoside, nucleotide and nucleic acid catabolic process	$2.1x10^{-2}$
Extracellular matrix organization	$2.2x10^{-2}$
Response to oxygen levels	$5.0x10^{-2}$

Some genes are down-regulated in cells exposed to both AZC and canavanine. This group of genes can provide new insights into the general response to low levels of mistranslation because AZC and canavanine have similar effects in protein misfolding and aggregation, despite their different chemical properties.

69 of the down-regulated genes were shared by cells exposed to AZC and canavanine (Figure 23). The list of these genes and respective fold change can be found in Appendix 2.

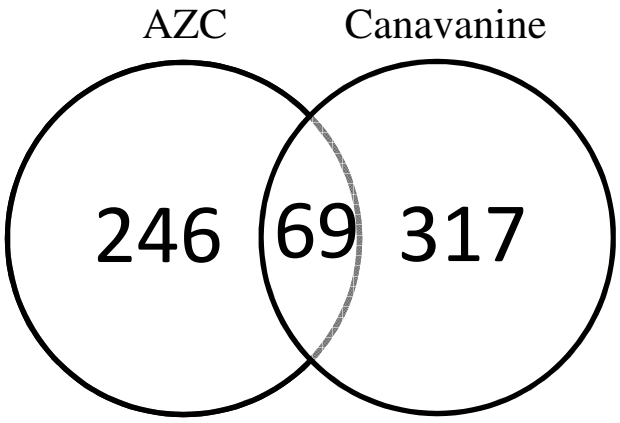


Figure 23. Genes down-regulated only with AZC, only with Canavanine and with both amino acid analogues.

Functional analysis of the 69 down-regulated genes, were performed using DAVID Bioinformatics Resources 6.7. GO annotation tool showed enrichment of genes that were mainly Membrane components (p-value=$2.3x10^{-2}$). Some of the genes annotated in this category are referred in the Table 3.

Table 3. Examples of genes that were annotated in the GO category Membrane components.

Genes	Proteins encoded
SLITRK6 RPRML TMEM80 TMEM129 TMEM107 PRRT2 SECTM1 STRA6	**Transmembrane proteins**
ANO7	**Protein involved in cell-cell interactions**
COL13A1 DHRS11	**Proteins related to cell matrix and cell adhesion**
FIBP	**Protein involved in tissue growth and differentiation**
FUT11	**Protein associated with protein modification**
PIGL	**Protein responsible for glycolipid biosynthesis**

Many of the down-regulated genes annotated in the cluster Membrane components codify transmembrane proteins. There are also genes that codify proteins involved in cell-cell interactions and cell matrix and cell adhesion, for instance collagen (COL13A1) and a secreted oxireductase (DRHS11). Cells exposed to amino acid analogues also down-regulated a gene responsible for mitogenesis and cell differentiation (FIBP). These genes (COL13A1, DRHS11 and FIBP) show similar fold changes in cells treated with AZC and with canavanine (Appendix 2).

From the 69 down-regulated genes, some of them were not annotated in any functional group. Table 4 shows these down-regulated genes.

Table 4. Examples of genes that were not annotated in any GO category but are also down-regulated with AZC and canavanine.

Genes	Proteins encoded
NES OBSL1 PDLIM7 CORO2A	**Proteins associated with cytoskeleton and intermediate filaments**
DDIT4 BAD	**Proteins responsible for inhibition of cell growth and stimulation of cell death**
TCOF1	**Protein involved in ribosomal DNA gene transcription**
RAB26	**Protein that regulate vesicular fusion and trafficking**
SLC25A15	**Mitochondria membrane proteins**
PEX26 ADLP	**Peroxisome membrane proteins**
TTC39A	**Protein that mediate protein-protein interactions**
PRODH	**Protein involved in proline degradation**

Again we can see that genes that codify many membrane proteins and organelles membrane were down-regulated. Genes that code for proteins associated with cytoskeleton and intermediate filaments, and regulate cellular trafficking were also down-regulated.

3.3.2 Functional analysis of up-regulated genes

Genes that were up-regulated by AZC are annotated mainly in the GO Terms: Transcription factor binding, Oxidation reduction and Response to unfolded proteins (Table 5).

Table 5. Functional categories of the genes up-regulated by AZC.

GO Annotation	p-value (Modified Fisher Exact p-value, EASE score)
Transcription factor binding	8.1×10^{-3}
Oxidation reduction	1.5×10^{-2}
Response to unfolded protein	2.1×10^{-2}
Transcription	2.7×10^{-2}
Transcription from RNA polymerase II promoter	3.8×10^{-2}
Response to protein stimulus	4.5×10^{-2}
RNA biosynthetic process	6.7×10^{-2}

The first GO category Transcription factor binding includes genes that codify transcription regulators and components of transcription factors. One example is the MAX gene that codifies a transcription regulator that may repress transcription via the recruitment of a chromatin remodeling complex. Two important categories of up-regulated genes were Response to unfolded protein and Response to protein stimulus. These categories included DDIT3 which is a transcription factor that is mainly activated by ER stress and promotes apoptosis; DNAJB1 which interacts with Hsp70 and stimulate its ATPase activity and HSPH1which codifies a heat shock protein that prevents the aggregation of denatured proteins in cells under severe stress and belongs to the Hsp70 family.

Once again cellular response to AZC and canavanine were similar (Table 6). For example, Regulation of transcription and Regulation of RNA metabolic processes were up-regulated in both cases. Regulation of transcription is mainly negative, and these categories include many transcription factors and components of transcription factors, namely zinc finger proteins. Some genes belong to the category Positive regulation of cell death, namely SAP30BP and CIDEC, which codify proteins that induce cell death by apoptosis. One gene which was not annotated in any category is AHSA1which codes for an activator of the Hsp90.

Table 6. Functional categories of the genes up-regulated by canavanine.

GO Annotation	**p-value (Modified Fisher Exact p-value, EASE score)**
Regulation of transcription, DNA-dependent	$1.3x10^{-6}$
Regulation of RNA metabolic process	$2.4x10^{-6}$
Regulation of transcription	$1.9x10^{-5}$
Negative regulation of gene expression	$5.5x10^{-4}$
Negative regulation of cellular biosynthetic process	$1.6x10^{-3}$
Transcription from RNA polymerase II promoter	$3.3x10^{-3}$
Positive regulation of cell death	$2.5x10^{-2}$

20 genes up-regulated in cells exposed to AZC were also up-regulated in cells exposed to canavanine (Figure 24).

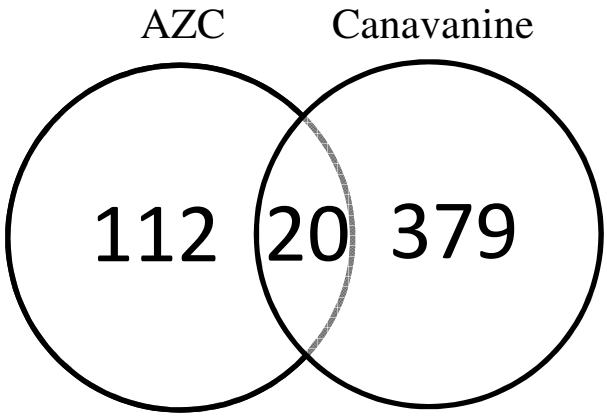


Figure 24. Genes up-regulated only with AZC, only with Canavanine and with both amino acid analogues.

Because of the low number of genes that were up-regulated by both AZC and canavanine, variable genes did not show enrichment for any particular functional group (Table 7).

Table 7. Examples of genes that were up-regulated by both AZC and canavanine

Genes	Proteins encoded
AFG3L1P BBS12	**Proteins involved in protein quality control**
SETD4 HOXA4 ZNF320	**Proteins involved in transcription process and regulation**
CD81	**Cell surface protein**
GUSBP1	**Beta-glucuronidase**
SRGAP2	**Protein involved in membrane protrusion**
SMPD3	**Proteins involved in cell stress response**

The list of the 20 up-regulated genes can be found in the Appendix 3. Some of these genes are related with the stress response to accumulation of unfolded proteins, and be involved in protein quality control. AFG3L1P codifies an ATP-dependent zinc metalloprotease, which seems to be involved in protein quality control, degrades non-assembled mitochondrial inner membrane proteins and can also function as a chaperone mediating the assembly of membrane-associated ATP-synthetase. The protein encoded by

BBS12 is a molecular chaperone that aids in protein folding upon ATP hydrolysis, and interacts with two other chaperonin-like BBS proteins, namely MKKS/BBS6 and BBS10 to form a chaperonin complex. Some genes are related to transcription process, such as SETD4, that probably functions by methylating histones on lysine residue, HOXA4 which codifies a DNA-binding transcription factor that may regulate gene expression, morphogenesis and differentiation and finally ZNF320, a gene that codifies the zinc finger protein 320. AZC and canavanine lead also to the up-regulation of SMPD3. This gene codifies a phosphodiesterase that hydrolyzes sphingomyelin to phosphocholine and ceramide. Ceramide is important in the coordination of eukaryotic stress responses and directs cells to the execution of specific programs of cell cycle arrest, differentiation and apoptosis.

Chapter IV – Discussion, Conclusions and Future Perspectives

4.1 Overview

The mRNA mistranslation phenomenon has been attracting increasing attention during the last few years. New discoveries associate mistranslation to generation of phenotypic diversity and evolution of the genetic code in some organisms, such as *Candida albicans.* New proteins formed can have advantageous functions, for example can allow the survival in toxic environments, creating a selective advantage (Miranda, Rocha et al. 2007; Moura, Paredes et al. 2010). But, mistranslation is also frequently associated to many diseases, in mice and humans (Nangle, Motta et al. 2006; Drummond and Wilke 2008; Schimmel 2008). Amino acid misincorporation into protein, due to loss of translational fidelity results in the formation of aberrant proteins whose accumulation usually leads to the formation of protein aggregates that are extremely toxic to cells and can lead to apoptosis (Schimmel 2008). The connection of mistranslation and human diseases makes this, an interesting subject in many scientific fields, for example, in biomedicine and in the pharmaceutical industry. Some of the toxicity mechanisms of misfolded proteins and protein aggregates have been elucidated in the past few years, but a comprehensive study of the magnitude of morphological cell alterations and cell stress response is essential to understand the effects of mistranslation and its connections with diseases.

To understand how human cells cope with error accumulation during protein synthesis, an *in vitro* model of mistranslation was used here. AZC and canavanine two amino acid analogues were used to induce mistranslation in human HEK293FT cells. These amino acid analogues are misincorporated into proteins instead of the natural amino acids and lead to the formation of misfolded proteins (Rodgers and Shiozawa 2008). With the objective of testing the level of toxicity caused by incorporation of these amino acid analogues a cellular viability assay was performed (MTT). Cells were also observed using light microscopy and analyzed by flow cytometry to detect the number of apoptotic cells for each drug concentration used. These analyses were important to determine levels of mistranslation that did not cause high level of cell death, which would introduce noise in our study of the cellular responses to the accumulation of misfolded proteins.

Studies with HEK cells expressing aggregation-prone proteins, such as huntingtin, have shown that protein aggregation leads to cell cycle arrest primarily at the G_2-M

boundary (Bence, Sampat et al. 2001). Cell cycle analyses were also carried out in this study to verify if this low level of mistranslation could lead to cell cycle arrest.

Human cells possess a variety of protein quality control mechanisms to minimize the accumulation of aberrant proteins and their toxic effects. Molecular chaperones are one of them (Bukau, Weissman et al. 2006). Chaperones are crucial for maintaining the native protein conformation and preventing nonspecific aggregation (Zhang, Beuron et al. 2002). When cells detect accumulation of misfolded proteins, they elevate the expression of molecular chaperones. The primary form of regulation is at the level of transcription (Morimoto 2008).

Another mechanism of protein quality control is the ubiquitin-proteasome system that degrades proteins by proteolysis. Misfolded proteins are in general polyubiquitinated prior to degradation by the 26S proteasome (Dahlmann 2007). Our data showed that proteins were conjugated with ubiquitin in cells exposed to AZC and canavanine.

Additionally, protein aggregation impair the function of the ubiquitin-proteasome system, which could lead to cellular deregulation and cell death (Bence, Sampat et al. 2001). In our cells proteasome activity did not change significantly suggesting that the aberrant proteins may accumulate instead of being degraded.

Finally, microarray analyses of the transcriptome showed that AZC and canavanine misincorporation deregulate gene expression, and highlighted the major protein functional categories affected.

4.2 Mistranslation has a negative impact on cellular viability and cell proliferation

It is well known that proteins with abnormal conformations can originate aggregates, which disrupt homeostasis and can lead to cell death. Especially in neurons, where protein aggregates have a strong negative impact in viability, as toxic aggregates and cannot be diluted by cell division. For this reason diseases associated to accumulation of misfolded proteins are mainly neurodegenerative disorders (Lee, Beebe et al. 2006). Cell death can be a consequence of the toxic effects of aberrant proteins, but can also be due to loss of function of essential proteins (Rochet 2006).

The MTT assay was used in our study to understand the effect of the amino acid analogues in cell viability and proliferation. This assay measures the metabolic activity of viable cells. Since proliferating cells are metabolically more active than non-proliferating cells, the assay is suitable not only for the determination of cell viability, but also for the determination of cell activation and proliferation (Wyllie 2009). Results showed a decrease in MTT reduction with increasing concentrations of the amino acid analogues, suggesting a negative impact of mistranslation in cellular viability and proliferation. In order to distinguish if this negative effect was related to an increase in cell death or decreased proliferation, nuclei were counted and flow cytometry analysis were carried out.

Nuclei stained with Hochest 33342 can be visualized by light microscopy which also allows differentiating death from viable cells. Our data showed that the percentage of apoptotic cells increased slightly only with increasing concentrations of AZC and canavanine. These results were confirmed by flow cytometry analysis with propidium iodide. Therefore, the range of concentrations of the amino acid analogues used in this work are likely to induce a low level of mistranslation and allow studing cellular responses to this phenomenon. Under these experimental conditions the low level of apoptotic cell death in HEK293FT cells may be due to the activation of PQC mechanisms which allow the cells to cope with the level of aberrant proteins formed.

Cell cycle analysis by flow cytometry of cells stained with propidium iodide showed decreased number of cells in G0-G1 phase, and increased number of cells in G2-M phase. AZC at 1.5mM decreased the number of cells in the S phase, but increased it at 2.5mM. Canavanine increased the percentage of cells in the S phase with increasing drug concentration.

Previous studies showed that canavanine is able to arrest cell cycle progression of human lung adenocarcinoma A549 cells in the G1 phase by inducing up-regulation of the negative cell cycle regulators p53 and $p21^{WAF1}$ (Ding, Matsukawa et al. 1999). Furthermore, Jurkat T cell cultures exposed to similar concentrations of canavanine (this amino acid analogue possess cytotoxicity to tumor cells), show higher percentage of cells in the S and G2-M phases before a remarkable increase in sub-G1 apoptotic cells. Apparently, 1.25mM and 2.5mM canavanine interrupts the completion of the S and M phase of Jurkat T cells. Some authors postulated that the effects of canavanine on the cell cycle may vary between cell types (Jang, Jun et al. 2002). The cell culture conditions, namely nutrients and gases availability can also fluctuate and may influence the state of proliferation of cells changing slightly the results between cell types, and even between cell replicates. Our results support those data. Canavanine at concentrations of 1.5mM and 2.5mM decreased the percentage of cells in G0-G1 and increased the percentage of cells in G2-M and S, suggesting that cell cycle progression to the S phase may be compromised and S phase completion may be interrupted.

Previous studies also refer the anti-tumor effects of AZC *in vitro* and showed a preferentially cell cycle arrest at the G2-M phase (Komander, Clague et al. 2009).

The only mechanism of toxicity reported so far to cause these effects in cell cultures involve the misincorporation of AZC and canavanine into proteins which may disrupt critical proteins required for the progression of the cell cycle. Apoptotic cell death increases slightly in the presence of these amino acid analogues and cells seem to arrest the progression of the cell cycle and be prevented from completing the phase S of DNA replication.

4.3 Mechanisms of Protein quality control in HEK293FT Cells

A brief revision of the primary responses to misfolded proteins is illustrated in Figure 25. Misfolded proteins that accumulate in cells can be refolded by molecular chaperones, degraded by UPS-mediated proteolysis, aggregate in the cytoplasm or in the nucleus or be secreted to the extracellular space (Ciechanover and Brundin 2003). In this study we aimed at testing whether misfolded proteins are targeted for degradation or whether the proteasome activity is compromised.

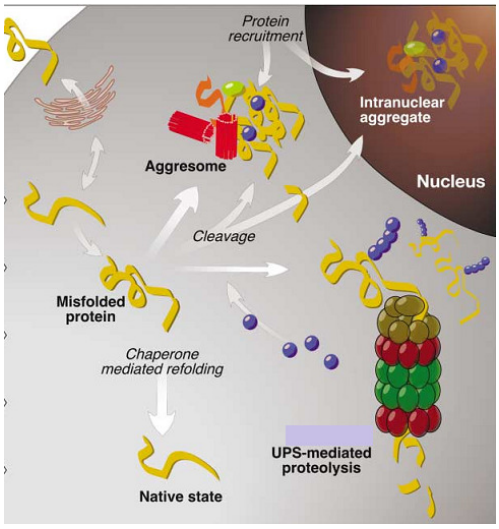


Figure 25. Primary responses to the accumulating of misfolded proteins. Misfolded proteins may be refolded by chaperones, may be targeted from degradation and may accumulate in aggregates in cytoplasm, nucleus or extracellular space, and sequester additional proteins [adapted from (Papa and Hochstrasser 1993)]

The accumulation of misfolded proteins triggered a stress response in HEK293FT cells. SDS-PAGE showed that AZC and canavanine increased ubiquitin conjugated proteins. These results are concordant with the study of Dasuri and colleagues who exposed neuronal and astrocyte cell cultures to AZC and canavanine and showed time-dependent increase of ubiquitinated proteins (Dasuri, Ebenezer et al. 2011).

It was reported that proteasome activity may be altered (increase or decrease) by accumulation of misfolded proteins. Parsell and Lindquist have shown that it is likely that an increase in damaged and misfolded proteins following heat shock up-regulate proteasome activity (Canu, Barbato et al. 2000). Consistent with this hypothesis, other

studies have demonstrated that proteasome activity can be rapidly elevated in response to mild oxidative stress (Reinheckel, Ullrich et al. 2000). This increase in proteasome activity seems to be due to accumulation of misfolded proteins up to a certain level. When the capacity of the proteasome to degrade proteins is overwhelmed, with rates of formation/delivery of misfolded proteins exceeding the capacity of the UPS to target and proteolytically degrade them, misfolded proteins can cause proteasomal stress and inhibit its activity (Papa and Hochstrasser 1993).

A number of studies refer UPS dysfunction in neurodegenerative disorders, such as Parkinson's and Alzheimer's diseases, caused by accumulation of aggregation-prone proteins (Papa and Hochstrasser 1993; Keller, Hanni et al. 2000). It was also demonstrated that soluble aggregated proteins inhibit the ubiquitin system (Bence, Sampat et al. 2001). In this study, Bence and colegues demonstrated in a cell cultured model that overexpression of mutant Huntingtin lead to proteasome inhibition. Another study showed that UPS inhibition only occurs under conditions of increased stress in cells expressing mutant Huntingtin (Amerik, Swaminathan et al. 1997). This increased proteotoxic stress is due to an elevated number of protein aggregates.

Our data indicates that AZC and canavanine do not alter proteasome activity likely because misfolded proteins are being efficiently targeted and degraded by the UPS and do not lead to proteasome stress. In other words, misincorporation of those amino acid analogues was not sufficient to alter proteasome activity. Higher concentrations of amino acid analogues may be required to alter proteasome activity. Possibly the cell type model can also influence the results, but this issue needs to be further investigated.

Since under the experimental conditions used in this work the accumulation of polyubiquitinated proteins was not a consequence of UPS impairment, they may accumulate despite the presence of a largely functional UPS. For example, if ubiquitinated substrates are sequestered into larger protein aggregates before reaching proteasome, this could explain the increase in ubiquitin conjugates observed in the SDS-PAGE. Ubiquitination can be involved in other processes such as clearance of aggregated proteins by macroautophagy which is a protective response that defends cells against the toxic insults of aggregation-prone proteins (Maynard, Böttcher et al. 2009). Deubiquitinating enzymes (DUBs) catalyze the removal of ubiquitin from proteins thus playing an important role in the editing of ubiquitination state of proteins and in the recycling of ubiquitin. Also these DUBs need to undergo structural rearrangements to adopt an active conformation

(Komander, Clague et al. 2009). If AZC and canavanine are being misincorporated in DUBs this can lead to inhibition of deubiquitination and consequently the increase in ubiquitinated proteins in cells. In yeast, disruption of DUBs leads to accumulation of polyubiquitn proteins and to depletion of cellular ubiquitin pools (Papa and Hochstrasser 1993; Amerik, Swaminathan et al. 1997; Canu, Barbato et al. 2000).

4.4 Gene transcription alterations

Many studies have reported several mechanisms that are activated in response to mistranslation and proteotoxic stress. However, a global view of the transcriptional mechanisms that are activated to counteract mistranslation in mammalian cells is still missing.

In *Saccharomyces cerevisiae* sub-lethal concentrations of AZC induced expression of heat-shock factor-regulated genes and selectively repressed the expression of ribosomal protein genes. Treatment with canavanine was less potent than with AZC and did not activated heat-shock factor (Trotter, Kao et al. 2002).

Human cells communicate with each other to respond to physiological stimulus and the diversity of human cell types may be accompanied by a corresponding diversity in their response to stress. Also, these cells have many physiological behaviors that contribute to their stress response, such as apoptosis and complex intercellular interactions, that are not present in yeast (Murray, Whitfield et al. 2004).

The HEK293FT cells exposed to amino acid analogues down-regulate genes mainly associated with the extracellular matrix which provides structural support to the cells and allows communication between them. Down-regulation of genes that codify proteins required for the maintenance of the extracellular matrix, namely collagen and laminins among others, transmembrane proteins and proteins involved in cell-cell interactions can compromise cell adhesion, and also communication. Additionally, cell morphology can be altered because of the down-regulation of several genes that codify proteins that maintain the cytoskeleton and intermediate filaments.

It would be interesting to carry out a detailed analysis of cell morphology and organization in culture to evaluate whether those gene expression alterations result in altered morphology. Since the extracellular matrix and controlled cell adhesion are essential for a coordinated morphogenesis and growth during embryonic development and since HEK293FT cells are embryonic kidney cells, it is reasonable to assume that AZC and canavanine will alter their morphology (Kuschel, Steuer et al. 2006).

Our microarray data also showed that genes involved in tissue growth and differentiation were down-regulated by AZC and canavanine suggesting that these processes are being compromised. The other main functional category down-regulated was the response to oxygen. This was a surprising result because it is known that accumulation

of misfolded proteins lead ER stress and consequently UPR activation which yields ROS accumulation (Haynes, Titus et al. 2004). In other words, we were expecting up-regulation of genes involved in oxygen metabolism.

Functional class analysis of genes up-regulated by AZC identified Transcription factor binding and Transcription as the main categories affected. Many transcription regulators were up-regulated, especially those that repress transcription, for example MAX. This suggests that overall transcription may be repressed in order to avoid synthesis of proteins that could misfold and compromise cell viability. Canavanine also lead to up-regulation of genes involved in Regulation of transcription and Negative regulation of gene expression. Additionally some genes up-regulated by canavanine participate in Positive regulation of cell death by apoptosis, which may indicate that misincorporation of canavanine does have a negative impact on viability. In viability assays we see an increase in the number of apoptotic cells, but this is very small, suggesting that this amino acid analogue, at these concentrations tested has low toxicity.

AZC exposure up-regulated genes involved in Protein Quality Control, Response to unfolded protein and Response to protein stimulus, namely DDIT3, DNAJB1 and HSPH1 which belong to the classic UPR pathway (Dombroski, Nayak et al. 2010). DDIT3 is a transcription factor that promotes apoptosis, DNAJB1 stimulates HSP70 ATPase activity and HSPH1is a heat shock protein that prevents aggregation of denatured proteins. The genes up-regulated by canavanine were not annotated in any of these functional classes, however the AHSA1, an activator of HSP90, which also belongs to the classic UPR pathway, was up-regulated (Dombroski, Nayak et al. 2010). Two genes, AFG3L1P and BBS12 that seem to be involved in protein quality control were also up-regulated by AZC and canavanine. AFG3L1P codifies a protein that can function as a chaperone and can degrade non-assembled mitochondrial inner membrane proteins, while BBS12 is a molecular chaperone. In spite of being involved in protein quality control these genes have not been reported to participate in UPR pathway.

Other up-regulated genes by AZC and canavanine are involved in transcription processes, negative transcription regulation and in cell stress response. For example, the SMPD3 gene which participates in the coordination of eukaryotic stress responses was identified in our data set. It is known that several stress agents, namely chemotherapeutic drugs and heat regulated sphingolipid metabolism by increasing the activity and concentration of enzymes that cause ceramide accumulation. Recent studies suggest that

ceramide is essential for mediating many of these stress responses. Ceramide can direct cells to cell cycle arrest, through retinoblastoma protein (Rb) activation and apoptosis, through caspases activation (Hannun and Luberto 2000; Nikolova-Karakashian and A.Rozenova 2010).

4.5 Conclusions and Future Perspectives

The objective of this study was to elucidate the cellular responses to mistranslation. For this, we have exposed human HEK293FT cells to the proline and arginine analogues, AZC and canavanine. The data demonstrate that cells counteract the negative effects of mistranslation by activating mechanisms that prevent progression of the cell cycle and the completion of the phase S. Apoptosis is also activated but to a lower extent.

As expected, HEK293FT cells activate mechanisms of protein quality control to counteract protein misfolding and aggregation induced by AZC and canavanine. Proteins are targeted for degradation by the ubiquitin-proteasome system, but the activity of the proteasome was not altered by the accumulation of misfolded proteins, suggesting that the level of misfolded proteins and aggregates was low. Despite this UPR was activated. Indeed, genes that belong to this classical pathway, namely DDIT3, DNAJB1, HSPH1 and AHSA1 were up-regulated to neutralize the accumulation of misfolded proteins and avoid protein aggregation. Other genes that are involved in protein quality control were also up-regulated, namely AFG3L1P and BBS12. We could complete this work assessing the levels of some of the heat shock proteins, for example Hsp70, Hsp40 and Hsp90 by western blot, because regulation of molecular chaperones can also be at translational level (Dasuri, Ebenezer et al. 2011).

As mentioned before, an unexpected result was the down-regulation of genes involved in the response to oxygen levels. It would be interesting to assess if ROS accumulate in these cells. Indeed, a recent study using AZC and canavanine to generate abnormal proteins in aging cells and neurodegeneration showed increased oxidative stress in both neurons and astrocytes (Dasuri, Ebenezer et al. 2011).

Our study shows therefore that mistranslation induced by amino acid analogues is a good model system to study the biology of protein misfolding and aggregation. It would be important to test additional drug concentrations and exposure time to obtain a more detailed understanding of the cellular responses to protein aggregation. It would also be interesting to evaluate protein aggregation in more detail using other techniques such as congo red or thioflavin staining. Finally, it would be exciting to create other models to induce controlled mistranslation in mammalian cells, for example by expressing misreading tRNAs. These models of mistranslation have the potential to help elucidate the biology of protein aggregation diseases.

References

Agris, P. F. (2004). "Decoding the genome: a modified view." Nucleic Acids Res **32**(1): 223-238.

ALLOZYNE. (2010, 10-12-2010). "ALLOZYNE Biociphering Platforms " Retrieved 28-01, 2011, from http://www.allozyne.com/platform.html.

Amerik, A. Y., S. Swaminathan, et al. (1997). "In vivo disassembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome." EMBO J **16**(16): 4826-4838.

Asahara, H., H. Himeno, et al. (1993). "Recognition Nucleotides of Escherichia coli tRNALeu and Its Elements Facilitating Discrimination from tRNASer and tRNATyr." Journal of Molecular Biology **231**(2): 219-229.

Baldwin, A. N. and P. Berg (1966). "Transfer ribonucleic acid-induced hydrolysis of valyladenylate bound to isoleucyl ribonucleic acid synthetase." J. Biol. Chem.(241): 839–845.

Bence, N. F., R. M. Sampat, et al. (2001). "Impairment of the Ubiquitin-Proteasome System by Protein Aggregation." Science **292**: 1552-1555.

Bukau, B., J. Weissman, et al. (2006). "Molecular Chaperones and Protein Quality Control." Cell **125**(3): 443-451.

Canu, N., C. Barbato, et al. (2000). "Proteasome Involvement and Accumulation of Ubiquitinated Proteins in Cerebellar Granule Neurons Undergoing Apoptosis." The Journal of Neuroscience **20**(2): 589-599.

Ciechanover, A. and P. Brundin (2003). "The Ubiquitin Proteasome System in Neurodegenerative Diseases: Sometimes the Chicken, Sometimes the Egg." Neuron **40**(2): 427-446.

Cole, C., A. Sobala, et al. (2009). "Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs." RNA(15): 2147–2160.

Cuervo, A. M. (2004). "Autophagy: in sickness and in health." TRENDS in Cell Biology **14**(2): 70-77.

Dahlmann, B. (2007). "Role of proteasomes in disease." BMC Biochemistry **8**(Suppl 1): S3.

Dantuma, N. P. and K. Lindsten (2010). "Stressing the ubiquitin-proteasome system." Cardiovascular Research (85): 263-271.

Dasuri, K., P. J. Ebenezer, et al. (2011). "Amino acid analogue toxicity in primary rat neuronal and astrocyte cultures: implications for protein misfolding and TDP-43 regulation " J Neurosci Res. **89**(9): 1471–1477.

Ding, Y., Y. Matsukawa, et al. (1999). "Growth inhibition of A549 human lung adenocarcinoma cells by L-canavanine is associated with p21/WAF1 induction." Jpn. J. Cancer Res.(90): 69–74.

Dobson, C. M. (2004). "Principles of protein folding, misfolding and aggregation." Semin Cell Dev Biol **15**(1): 3-16.

Dombroski, B. A., R. R. Nayak, et al. (2010). "Gene Expression and Genetic Variation in Response to Endoplasmic Reticulum Stress in Human Cells." The American Journal of Human Genetics **86**(5): 719-729.

Drummond, D. A. and C. O. Wilke (2008). "Mistranslation-Induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution." Cell (134): 341–352.

Ecroyd, H. and J. A. Carver (2008). "Unraveling the mysteries of protein folding and misfolding." IUBMB Life **60**(12): 769-774.

Eriani, G., M. Delarue, et al. (1990). "Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs." Nature (347): 203–206.

Farabaugh, P. J. and G. R. Bjork (1999). "How translational accuracy influences reading frame maintenance." EMBO J **18**(6): 1427-1434.

Favorova, O. (1984). "Superspecificity of aminoacyl-tRNA synthetases." Mol Biol **18**(1): 205-226.

Florentz, C. and M. Sissler (2001). "Disease-related versus polymorphic mutations in human mitochondrial tRNAs. Where is the difference?" EMBO Rep **2**(6): 481-486.

Gallant, J. and L. Palmer (1979). "Error propagation in viable cells " Mechanisms of Ageing and Development **10** 27-38

Gomes, A., I. Miranda, et al. (2007). "A genetic code alteration generates a proteome of high diversity in the human pathogen Candida albicans." Genome Biology **8**(10): R206.

Gregersen, N., P. Bross, et al. (2006). "Protein misfolding and human disease." Annu Rev Genomics Hum Genet **7**: 103-124.

Hannun, Y. A. and C. Luberto (2000). "Ceramide in the eukaryotic stress response." Trends in Cell Biology **10**(2): 73-80.

Hausmann, C. D. and M. Ibba (2008). "Aminoacyl-tRNA synthetase complexes:molecular multitasking revealed." FEMS Microbiol Rev(32): 705–721.

Haynes, C. M., E. A. Titus, et al. (2004). "Degradation of Misfolded Proteins Prevents ER-Derived Oxidative Stress and Cell Death." Molecular Cell **15**: 767–776.

Hendrickson, T. L., V. d. Crécy-Lagard, et al. (2004). "Incorporation of nonnatural amino acids into proteins " Annu. Rev. Biochem.(73): 147–176.

Herczenik, E. and M. F. Gebbink (2008). "Molecular and cellular aspects of protein misfolding and disease." FASEB J **22**(7): 2115-2133.

Hopper, A. K., D. A. Pai, et al. (2010). "Cellular dynamics of tRNAs and their genes." FEBSLetters **584**: 310-317.

Huang, D. W., B. T. Sherman, et al. (2008). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nat. Protocols **4**(1): 44-57.

Huang, D. W., B. T. Sherman, et al. (2009). "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists." Nucleic Acids Research **37**(1): 1-13.

Igloi, G. L. and E. Schiefermayr (2009). "Amino acid discrimination by arginyl-tRNA synthetases as revealed by an examination of natural specificity variants." FEBS Journal **276** 1307–1318.

Ishitani, R., S. Yokoyama, et al. (2008). "Structure, dynamics, and function of RNA modification enzymes." Current Opinion in Structural Biology (18): 330–339.

Jang, M. H., D. Y. Jun, et al. (2002). "Arginine antimetabolite L-canavanine induces apoptotic cell death in human Jurkat T cells via caspase-3 activation regulated by Bcl-2 or Bcl-xL." Biochemical and Biophysical Research Communications **295** 283-288.

Jhon, J. S. and Y. K. Kang (2007). "Conformational Preferences of Proline Analogues with Different Ring Size." J. Phys. Chem. **111**: 3496-3507.

Johnston, J. A., C. L. Ward, et al. (1998). "Aggresomes: A Cellular Response to Misfolded Proteins." The Journal of Cell Biology **143**(7): 1883-1898.

Kapp, L. and K. Lorsch (2004). "The Molecular Mechanics of Eukaryotic Translation " Annual Reviews of Biochemystry (73): 657-704.

Keller, J. N., K. B. Hanni, et al. (2000). "Impaired Proteasome Function in Alzheimer's Disease." Journal of Neurochemistry **75**(1): 436-439.

Komander, D., M. J. Clague, et al. (2009). "Breaking the chains: structure and function of the deubiquitinases." Nat Rev Mol Cell Biol **10**(8): 550-563.

Kuschel, C., H. Steuer, et al. (2006). "Cell adhesion profiling using extracellular matrix protein microarrays." BioTechniques **40**: 523-531.

Larsen, K. E. and D. Sulzer (2002). "Autophagy in neurons: a review." Histol. Histopathol.(17): 897–908.

Lee, J. W., K. Beebe, et al. (2006). "Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration." Nature **443**: 50-55.

Lenhard, B., O. Orellana, et al. (1999). "tRNA recognition and evolution of determinants in seryl-tRNA synthesis." Nucleic Acids Research **27**(3): 721-729.

Levine, L. R., L. Mosoni, et al. (1996). "Methionine residues as endogenous antioxidants in proteins " Proc. Natl Acad. Sci. USA(93): 15036-15040.

Levinger, L., M. Morl, et al. (2004). "Mitochondrial tRNA 3' end metabolism and human disease." Nucleic Acids Res **32**(18): 5430-5441.

Lewin, B. (2004). Genes VIII.

Ling, J., N. Reynolds, et al. (2009). "Aminoacyl-tRNA Synthesis and Translational Quality Control." Annu. Rev. Microbiol.(63): 61–78.

Lodish, H., A. Berk, et al. (2005). Molecular Cell Biology New York, W. H. Freeman and Company.

Martinis, S. A. and M. T. Boniecki (2009). "The balance between pre- and post-transfer editing in tRNA synthetases." FEBS Letters **584**: 455–459.

Maynard, C. J., C. Böttcher, et al. (2009). "Accumulation of ubiquitin conjugates in a polyglutamine disease model occurs without global ubiquitin/proteasome system impairment." Proceedings of the National Academy of Sciences **106**(33): 13986-13991.

Miranda, I., R. Rocha, et al. (2007). "A Genetic Code Alteration Is a Phenotype Diversity Generator in the Human Pathogen Candida albicans." PLoS one(10): 1-11.

Miranda, I., R. Silva, et al. (2006). "Evolution of the genetic code in yeasts " Yeast (23): 203-213.

Miyoshi, K., T. Miyoshi, et al. (2010). "Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production." Mol Genet Genomics(2): 95-103.

Morimoto, R. I. (2008). "Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging." Genes Dev. **22**(11): 1427-1438.

Moura, G. R., L. C. Carreto, et al. (2009). "Genetic code ambiguity: an unexpected source of proteome innovation and phenotypic diversity." Current Opinion in Microbiology (12): 1-7.

Moura, G. R., J. A. Paredes, et al. (2010). "Development of the genetic code: Insights from a fungal codon reassignment." FEBS Letters(584): 334-341.

Muchowski, P. J. and J. L. Wacker (2005). "Modulation of neurodegeneration by molecular chaperones." Nature Reviews(6): 11-22.

Murray, J. I., M. L. Whitfield, et al. (2004). "Diverse and specific gene expression responses to stresses in cultured human cells " Molecular Biology of the Cell **14**: 2361-2374.

Nakanishi, K. and O. Nureki (2005). "Recent Progress of Structural Biology of tRNA Processing and Modification " Mol. Cells **19**(2): 157-166

Nangle, L. A., V. d. C. Lagard, et al. (2002). "Genetic Code Ambiguity - Cell viability related to the severity of editing defects in mutant tRNA synthetases " THE JOURNAL OF BIOLOGICAL CHEMISTRY **277**(48): 45729–45733.

Nangle, L. A., C. M. Motta, et al. (2006). "Global Effects of Mistranslation from an Editing Defect in Mammalian Cells " Chemistry & Biology(13): 1091–1100.

Nikolova-Karakashian, M. N. and K. A.Rozenova (2010). "Ceramide in stress response." Adv Exp Med Biol. **688**: 86-108.

O'Donoghue, P. and Z. Luthey-Schulten (2003). "On the Evolution of Structure in Aminoacyl-tRNA Synthetases." MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS **67**(4): 550–573.

Papa, F. R. and M. Hochstrasser (1993). "The yeast DOA4 gene encodes a deubiquitinating enzyme related to a product of the human tre-2 oncogene." Nature **366**(6453): 313-319.

Park, S. G., P. Schimmel, et al. (2008). "Aminoacyl tRNA synthetases and their connections to disease." PNAS **105**(32): 11043–11049.

Pines, J. and C. Lindon (2005). "Proteolysis: anytime, any place, anywhere?" Nat Cell Biol(7): 731–735.

Pouplana, L. R. d. and P. Schimmel (2001a). "Aminoacyl-tRNA synthetases: potential markers of genetic code development." TRENDS in Biochemical Sciences **26**(10): 591-596

Pouplana, L. R. d. and P. Schimmel (2001b). "Two Classes of tRNA Synthetases Suggested by Sterically Compatible Dockings on tRNA Acceptor Stem." Cell **104**: 191-193.

Powers, E. T. and W. E. Balch (2008). "Costly Mistakes: Translational Infidelity and Protein Homeostasis." Cell (134): 204-206.

Raam, B. J. v. and G. S. Salvesen (2010). "Transferring Death: A Role for tRNA in Apoptosis Regulation." Molecular Cell **37**(5): 591-592.

Rabilloud, T., J.-M. Strub, et al. (2002). "Comparative Proteomics as a New Tool for Exploring Human Mitochondrial tRNA Disorders." Biochemistry **41**: 144-150.

Reinheckel, T., O. Ullrich, et al. (2000). "Differential Impairment of 20S and 26S Proteasome Activities in Human Hematopoietic K562 Cells during Oxidative Stress." Archives of Biochemistry and Biophysics **377**(1): 65-68.

Reynolds, N. M., B. A. Lazazzera, et al. (2010). "Cellular mechanisms that control mistranslation " Nature Reviews **8**: 849-856.

Richmond, M. H. (1962). "The effect of amino acid analogues on growth and protein synthesis in microorganisms " Bacteriol Rev. **26**(4): 398–420.

Rochet, J.-C. (2006). "Errors in Translation Cause Selective Neurodegeneration." ACS Chemical Biology **1**(9): 562-566.

Rodgers, K. J. and N. Shiozawa (2008). "Misincorporation of amino acids analogues into proteins by biosynthesis " The International Journal of Biochemistry and Cell Biology (40): 1452-1466.

Saeed, A., N. K. Bhagabati, et al. (2006). "TM4 microarray software suite." Methods Enzymol(411): 134-193.

Santos, A. M., C. Cheesman, et al. (1999). "Selective advantages created by codon ambiguity allowed for the evolution of an alternative genetic code in Candida Spp." Mol. Micobiol. (31): 937-947.

Scheper, G., M. Knaap, et al. (2007). "Translation matters: protein synthesis defects in inherited disease." Nature Reviwes **8**: 711-723.

Schimmel, P. (2008). "Development of tRNA synthetases and connection to genetic code and disease " Protein Science(17): 1643-1652.

Sekine, S., O. Nureki, et al. (1999). "The identity determinants required for the discrimination between tRNAGlu and tRNAAsp by glutamyl-tRNA synthetase from Escherichia coli." Eur. J. Biochem. (261): 354-360.

Sherman, J. M. and D. Söll (1996). "Aminoacyl-tRNA Synthetases Optimize Both Cognate tRNA Recognition and Discrimination against Noncognate tRNAs." Biochemistry **35**(2): 601-607.

Simon, R., E. L. Korn, et al. (2003a). "Design and Analysis of DNA Microarray Investigations." Springer verlag.

Sprinzl, M. and F. Cramer (1975). "Site of aminoacylation of tRNAs from Escherichia coli with respect to the 2'- or 3'-hydroxyl group of the terminal adenosine." Biochemistry **72**(8): 3049-3053.

Trotter, E. W., C. M.-F. Kao, et al. (2002). "Misfolded Proteins Are Competent to Mediate a Subset of the Responses to Heat Shock in Saccharomyces cerevisiae." THE JOURNAL OF BIOLOGICAL CHEMISTRY **277**(47): 44817–44825.

Voges, D., P. Zwickl, et al. (1999). "The 26S proteasome: a molecular machine designed for controlled proteolysis " Annu Rev Biochem(68): 1015-1068.

Wacker, J. L., M. H. Zareie, et al. (2004). "Hsp70 and Hsp40 atenuate formation os spherical and annular polyglutamine oligomers by partitioning monomer." Nature Struct. Mol. Biol. (11): 1215-1222.

Wills, N. M. and J. F. Atkins (2006). "The potential role of ribosomal frameshifting in generating aberrant proteins implicated in neurodegenerative diseases." RNA **12**.

Worthen, D. R., L. Chien, et al. (1998). "L-Canavanine modulates cellular growth, chemosensitivity and P-glycoprotein substrate accumulation in cultured human tumor cell lines." Cancer Letters (132): 229-239.

Wyllie, A. H. (2009). Apoptosis, Cell Death, and Cell Proliferation Manual, Roche Applied Science 74-84.

Xie, W., L. A. Nangle, et al. (2007). "Long-range structural effects of a Charcot–Marie–Tooth disease-causing mutation in human glycyl-tRNA synthetase." PNAS **104** (24): 9976–9981.

Xu, C., B. Bailly-Maitre, et al. (2005). "Endoplasmic reticulum stress: cell life and death decisions." The Journal of Clinical Investigation **115**(10): 2656-2664

Zhang, X., F. Beuron, et al. (2002). "Machinery of protein folding and unfolding." Current Opinion in Structural Biology(12): 231–238.

Appendix

1. Solutions

PBS

- 137 mM NaCl
- 2,7 mM KCl
- 10 mM Na_2HPO_4
- 1,8 mM KH_2PO_4

- pH was adjusted to 7.4

Lysis buffer for Proteasome activity detection

- 1mM EDTA
- 10 mM tris-HCl, pH 7.5
- 20% glycerol
- 4 mM DTT (dithiothreitol)
- 2 mM ATP

Proteasome Activity buffer

- 0,5 mM EDTA
- 50 mM Tris-HCl, pH 8
- 2 mM ATP
- 50 µM Suc-LLVY-MCA
- 10 µM MG132 (just in one of the duplicates)

Lysis buffer for Western Blot

- 50mM HEPES
- 150mM NaCl
- 2mM EGTA
- 1mM EDTA
- 1mM DTT
- 2mM Na_3VO_4
- 50mM NaF
- 1% Triton X-100

- pH was adjusted to 7.4

- supplemented with 1mM PMSF and 1µg/ml CLAP in the moment of lysis

Denaturating buffer for Western Blot

- 0,25 mM Tris, pH 6,8
- 4% SDS
- 200 mM DTT
- 20% glycerol
- 0,01% bromophenol blue

Running gel for Western Blot (10%)

H_2O	3,95 mL
Tris 1,5 M, pH 8,8	3,35 mL
40% Acrilamide (Acrilamide/Bisacrilamide solution, 37,5:1)	2,5 mL
20% SDS	100 µL
10% AMPS (0,1g in 1mL H_2O)	100 µL
TEMED	5 µL

Stacking gel for Western Blot (4%)

H_2O	3,1 mL
Tris 0,625 M, pH 6,8	125 mL
40% Acrilamide (Acrilamide/Bisacrilamide solution, 37,5:1)	0,5 mL
20% SDS	50 µL
10% AMPS (0,1g in 1mL H_2O)	50 µL
TEMED	5 µL

Electrophoresis buffer for Western Blot

- 100 mM Tris
- 100 mM Bicine
- 0,1% SDS

Electrotransfer buffer

TBS 10x

	2 L
200 mM Tris	24,2 g
1,37 M NaCl	80,0 g

- Adjust the pH to 7,6

TBS-T with 0,5% milk

	2 L
TBS 10x	200 mL
0,1% Tween	2 mL
0,5% milk	10 g
H_2O mQ	To add up to 2 L

2. List of genes down-regulated in cells exposed to AZC or to canavanine

Gene name	Gene symbol	AZC Fold Change(*)	Canavanine Fold Change(*)
ADAM metallopeptidase with thrombospondin type 1 motif, 4	ADAMTS4	0.78	0.50
anoctamin 7	ANO7	0.71	0.45
arachidonate 5-lipoxygenase	ALOX5	0.59	0.33
ArfGAP with GTPase domain, ankyrin repeat and PH domain 2	AGAP2	0.72	0.60
arrestin, beta 1	ARRB1	0.81	0.56
ATG16 autophagy related 16-like 2	ATG16L2	0.69	0.57
ATP-binding cassette sub-family D member 1 (Adrenoleukodystrophy protein)	ALDP	0.73	0.70
Bcl2 antagonist of cell death	BAD	0.71	0.73
beta-1,4-N-acetyl-galactosaminyl transferase 1	B4GALNT1	0.56	0.21
chemokine (C-C motif) receptor 10	CCR10	0.70	0.48
chromosome 6 open reading frame 108	C6orf108	0.77	0.76
chromosome 6 open reading frame 126	C6orf126	0.67	0.45
collagen, type XIII, alpha 1	COL13A1	0.65	0.56
coronin, actin binding protein, 2A	CORO2A	0.85	0.70
crystallin, mu	CRYM	0.72	0.48
cyclin I family, member 2	CCNI2	0.80	0.44
DAN domain family, member 5	DAND5	0.65	0.361
Dehydrogenase/reductase SDR family member 11	DHRS11	0.70	0.76
DNA-damage-inducible transcript 4	DDIT4	0.63	0.27
ectonucleoside triphosphate diphosphohydrolase 8	ENTPD8	0.80	0.48
exocyst -like pseudogene	LOC644145	0.65	0.36
family with sequence similarity 27, member A	FAM27A	0.79	0.45
fibroblast growth factor (acidic) intracellular binding protein	FIBP	0.73	0.71

fucosyltransferase 11 (alpha (1,3) fucosyltransferase)	FUT11	0.82	0.42
GNAS antisense RNA (non-protein coding)	GNASAS	0.71	0.58
hypothetical LOC100128163	LOC100128163	0.84	0.73
hypothetical LOC100130015	LOC100130015	0.65	0.41
hypothetical LOC100130948	LOC100130943	0.74	0.68
hypothetical LOC100134361	LOC100134361	0.64	0.62
hypothetical protein LOC729683	LOC729683	0.73	0.44
hypothetical protein MGC16121	MGC16121	0.75	0.42
intercellular adhesion molecule 4 (Landsteiner-Wiener blood group)	ICAM4	0.68	0.52
jumonji domain containing 7	JMJD7	0.75	0.56
Kazal-type serine peptidase inhibitor domain 1	KAZALD1	0.58	0.21
kelch domain containing 9	KLHDC9	0.79	0.53
leucine rich repeat and Ig domain containing 3	LINGO3	0.54	062
magnesium-dependent phosphatase 1	MDP1	0.85	0.75
NCK interacting protein with SH3 domain	NCKIPSD	0.77	0.60
nestin	NES	0.75	0.46
neurofilament, medium polypeptide	NEFM	0.67	0.36
obscurin-like 1	OBSL1	0.70	0.51
PDZ and LIM domain 7 (enigma)	PDLIM7	0.87	0.68
peroxisomal biogenesis factor 26	PEX26	0.73	0.53
phosphatidylinositol glycan anchor biosynthesis, class L	PIGL	0.75	0.44
polyamine oxidase (exo-N4-amino)	PAOX	0.70	0.46
potassium inwardly-rectifying channel, subfamily J, member 12	KCNJ12	0.71	0.69
proline dehydrogenase (oxidase) 1	PRODH	0.86	0.68
proline-rich transmembrane protein 2	PRRT2	0.72	0.51
Putative uncharacterized protein ENSP00000385341		0.75	0.72

pyroglutamylated RFamide peptide receptor	QRFPR	0.83	0.69
RAB26, member RAS oncogene family	RAB26	0.43	0.44
reprimo-like	RPRML	0.59	0.57
secreted and transmembrane 1	SECTM1	0.70	0.45
similar to hCG1804255	LOC100129794	0.88	0.62
similar to hCG2039305	LOC729041	0.83	0.63
SLIT and NTRK-like family, member 6	SLITRK6	0.60	0.39
solute carrier family 22 (organic anion/urate transporter), member 11	SLC22A11	0.63	0.17
solute carrier family 25 (mitochondrial carrier; ornithine transporter) member 15	SLC25A15	0.84	0.61
stimulated by retinoic acid gene 6 homolog (mouse)	STRA6	0.69	0.37
sulfotransferase family, cytosolic, 2A, dehydroepiandrosterone (DHEA)	SULT2A1	0.66	0.59
tetratricopeptide repeat domain 39A	TTC39A	0.67	0.72
transmembrane protein 107	TMEM107	0.86	0.68
transmembrane protein 129	TMEM129	0.62	0.60
transmembrane protein 80	TMEM80	0.69	0.76
Treacher Collins-Franceschetti syndrome 1	TCOF1	0.64	0.38
tripartite motif-containing 36	TRIM36	0.76	0.70
Uncharacterized protein ENSP00000382042		0.84	0.19
Yip1 domain family, member 2	YIPF2	0.76	0.61
zinc finger protein 780A	ZNF780A	0.78	0.62

* Fold change: below 1 – down-regulated genes; above 1 – up-regulated genes

3. List of genes up-regulated in cells exposed to AZC or to canavanine

Gene name	Gene symbol	AZC Fold Change(*)	Canavanine Fold Change(*)
AFG3 ATPase family gene 3-like 1 (S. cerevisiae)	AFG3L1	1.24	1.46
Bardet-Biedl syndrome 12	BBS12	1.21	1.32
CD81 molecule	CD81	1.14	1.11
cDNA FLJ23879 fis, clone LNG13743		1.38	2.05
CDNA FLJ25625 fis, clone STM02974		1.24	1.44
cDNA FLJ37880 fis, clone BRSTN2000220		1.28	2.20
chromosome 1 open reading frame 198	C1orf198	1.19	1.45
collagen-like tail subunit (single strand of homotrimer) of asymmetric acetylcholinesterase	COLQ	1.54	1.47
FSHD region gene 1 family, member B	FRG1B	1.17	1.28
glucuronidase, beta pseudogene	GUSBP1	1.26	1.71
guanylate cyclase 1, soluble, alpha 3	GUCY1A3	1.24	1.46
homeobox A4	HOXA4	1.10	1.58
melanoregulin	MREG	1.22	1.30
mitochondrial ribosomal protein L42 pseudogene 5	MRPL42P5	1.27	1.44
paired box 6	PAX6	1.17	1.57
SET domain containing 4	SETD4	1.31	1.69
SLIT-ROBO Rho GTPase activating protein 2	SRGAP2	1.17	1.80
sphingomyelin phosphodiesterase 3, neutral membrane (neutral sphingomyelinase II)	SMPD3	1.42	1.50
tuftelin 1	TUFT1	1.32	1.44
zinc finger protein 320	ZNF320	1.20	1.63

* Fold change: below 1 – down-regulated genes; above 1 – up-regulated genes